

RESU 3
B74739 standard: 147 AA
B74739
R74739
21-AN-1966 (first entry)
Rat ubiquitous nuclear receptor protein.
Ubiquitous nuclear receptor; drug design; disease diagnosis; therapy.
Rattus rattus
W0951373-A1.
18-MAY-1995
08-NOV-1994 94W0-0512683.
10-NOV-1993 9305-0152003.
Liao S, Song C;
WPI: 1995-194092/25.
N-PSDB: Q88761.
New ubiquitous nuclear receptor - used to develop prods. for use in diagnostic, drug design and therapeutic applications
Claim 5: Page 144: 195pp; English.
This protein produced by recombinant expression of genomic DNA isolated from rat vagina interacts with the response elements and network of regulatory factors in the human estrogen and receptor subfamily. The products can be used in detection, diagnosis and screening assays for substances which interact with UNKS for use in diagnosis, drug design and therapeutic applications.
Sequence 443 AA

```

Query Match: 97.28% Score 23175, 18.16, Length 443
Host Local Similarity: 97.18%, Pred. No. 4,60-1949
Matches: 4439, Coverage: 99.96, Mismatches: 4, Indels: 0, Gaps: 0

0Y 1 MSSTPSSTLPPF MSSTPSSTSAVSTLKEEEDLPPF SESSSSALVATLEDEDP 60
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1b 1 MSSTPSSTLPIPPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 60
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
0Y 61 EPRPRKTRVEMLHILTRVETKRSKGRIMWLSLSSKSLTAKVWVSSKSLKSLSSG 120
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1b 58 EPRKTPPRPRKTRVEMLHILTRVETKRSKGRIMWLSLSSKSLTAKVWVSSKSLK 120
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
0Y 121 TCGAAAPWPKKTVCTPRPKFAKMTKGVAFKQSTKSGGQVGGGQSPSPSSAASS 180
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1b 118 TCGDAAPMTKTVCTPRPKFAKMTKGVAFKQSTKSGGQVGGGQSPSPSSAASS 180
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
0Y 181 GQAAAPSPSPSSQSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 240
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1b 178 GPAAAPSPSPSSQSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 240
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
0Y 241 AHHVSRNAGGPRVHTTFLATISVPTVTPKAKYGRNGQVGGPDVATIKASILEM 300
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1b 238 AHPVSRNAGGPRVHTTFLATISVPTVTPKAKYGRNGQVGGPDVATIKASILEM 300
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
0Y 301 TAKKYNHEECITFLKDPVYSKQHPKAGVGPVNIIEFSSPMRGLDIAEVALIA 360
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1b 298 TAKKYNHEECITFLKDPVYSKQHPKAGVGPVNIIEFSSPMRGLDIAEVALIA 360
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
0b 298 TAKKYNHEECITFLKDPVYSKQHPKAGVGPVNIIEFSSPMRGLDIAEVALIA 360

```

09	461	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	42
10	468	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	47
11	471	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	52
12	474	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	57
13	477	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	62
14	480	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	67
15	483	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	72
16	486	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	77
17	489	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	82
18	492	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	87
19	495	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	92
20	498	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	97
21	501	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	102
22	504	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	107
23	507	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	112
24	510	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	117
25	513	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	122
26	516	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	127
27	519	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	132
28	522	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	137
29	525	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	142
30	528	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	147
31	531	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	152
32	534	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	157
33	537	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	162
34	540	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	167
35	543	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	172
36	546	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	177
37	549	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	182
38	552	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	187
39	555	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	192
40	558	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	197
41	561	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	202
42	564	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	207
43	567	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	212
44	570	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	217
45	573	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	222
46	576	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	227
47	579	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	232
48	582	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	237
49	585	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	242
50	588	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	247
51	591	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	252
52	594	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	257
53	597	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	262
54	600	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	267
55	603	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	272
56	606	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	277
57	609	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	282
58	612	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	287
59	615	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	292
60	618	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	297
61	621	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	302
62	624	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	307
63	627	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	312
64	630	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	317
65	633	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	322
66	636	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	327
67	639	INSEKAPNUPESPVALPOLVENEI LESTERKOPVALPEPOMEKXVSLKTISSVH	332
68	642	INSEKAPNUPESPVALPOL	

RESULT 4
W25035
ID W25035 Standard; Protein; 443 AA

15-000-1997 (first entry)

DE Rat ubiquitous nuclear receptor polypeptide

KM Ubiquitous nuclear receptor polypeptide; DNA binding domain;
KM therapy; drug design; diagnosis.

(5) *Rattus rattus*.

US5639616-A.

PD 17-JUN-1997.

PF 10-NOV-1993; 93115-0152003

PR 18-NOV-1994; 94US-0442411

[illegible]

XX

XX

N-PSIH; T79635.

PT RNA encoding ubiquitin

PI producing recombinant poly:peptide(s) and detecting RNA

PS C:\ADMIN\1; C:\ADMIN\77-80; 67pp; END; 157.

W250) shows the amino acid sequence of a rat ubiquitous nuclear receptor polypeptide (NR). This is useful in assays designed to

selected substances which interact with the DR which may potentially

design.

443 AA;

Quality Match	97.28	Score	2247.5	DB	18	Length	443
Exon Intron Similarity	37.18	Prod. No.	4.60-193				
Matches	433	Conservative	6	Mismatches	4	Indels	3
						Gaps	1

1 MSSPSSIDTIVPQNSQ₄TSAS₁IKK₂Q₁DP₁QSE₂SSAY₁V₁LPED₁ 60

bb 1 mssptssldlplpqngspqpslststspikeevqetdpppqseqssayiv---epedep '57

[illegible]

bb 58 ckrkkpppljhe lcrccqjhct bnoo lsogck fte rsnbhq qnyacrsd : 17

186

10b 118 tcmdmrfkq|cr|rkobvobvov|lsouj|lbb|kb|ppppptepaqss 177

181 CBPAASPGTSPASSGQSGHGGCTCTAACTMCTGAACTGCKPSPSPKVTWPICT 240

178 arbastrat eassssqqsqeeq[ql]a nrlm[nr]vav]gokrsfcdnkvtpw]d 037

QY 041 ADEGSEAWAGPFAHRRPAAISV,ETGFAKVGDTLLEEFVVALFASTETEMLEQ 390
 DB 248 ADEGSEAWAGPFAHRRPAAISV,ETGFAKVGDTLLEEFVVALFASTETEMLEQ 390
 QY 401 TARRNHDETECTEETFEYTSKQEFPAALQVEFINDPFEFSPAMEPLDNEVALIA 360
 DB 298 TARRNHDETECTEETFEYTSKQEFPAALQVEFINDPFEFSPAMEPLDNEVALIA 360
 QY 041 TARRNHDETECTEETFEYTSKQEFPAALQVEFINDPFEFSPAMEPLDNEVALIA 420
 DB 458 TARRNHDETECTEETFEYTSKQEFPAALQVEFINDPFEFSPAMEPLDNEVALIA 420
 QY 421 SEGVAFLEGEPEFPELSEIWEWE 445
 DB 418 SEGVAFLEGEPEFPELSEIWEWE 445
 RESULT 5
 R74738
 ID R74738 standard; Protein: 460 AA.
 XX
 AF R74738;
 XX 20 JAN-1996 (first entry)
 DE Human ubiquitons nuclear receptor protein.
 XX Ubiquitons nuclear receptor; drug design; disease diagnosis;
 KW therapy.
 XX Homo sapiens.
 OS
 PN W09513373-A1.
 XX
 PD 18-MAY-1995.
 XX
 PF 08-NOV-1994; 94MO-0812884.
 XX
 PK 10-NOV-1994; 9408-0192004.
 XX
 PA (ARCH) ARCH DEV CORP.
 XX
 P1 Liao S, Song C;
 XX
 DR WPI: 1995-194092/25.
 DR N-PSDB: Q88760.
 XX
 PT New ubiquitons nuclear receptor - used to develop prods. for use in
 PT diagnostic, drug design and therapeutic applications
 PS Claim 5; Page 140; 19pp; English.
 CC This protein produced by recombinant expression of genomic DNA
 CC isolated from human prostate cancer PC-3 cells interacts with
 CC the response elements and network of receptors in the thyroid
 CC hormone/reinoic acid receptor subfamily. The products can be used
 CC in detection, diagnosis and in screening assays for substances which
 CC interact with UMRs for use in diagnosis, drug design and
 CC therapeutic applications.
 CC
 SU Sequence 460 AA;
 QY Query Match 88.4%; Score 2044; DB 16; Length 460;
 Best Local Similarity 87.4%; Pred. No. 8,1e-175;
 Matches 402; Conservative 17; Mismatches 27; Indels 14; Gaps 4;
 QY 1 MSHPTSSDPTTVVWNGSGGQSLSAISPTKKEDEG-----LDPFPGSSNASSAVV V 52
 DB 1 MSHPTSSDPTTVVWNGSGGQSLSAISPTKKEDEG-----LDPFPGSSNASSAVV V 52
 QY 53 HLEPEHERRKKRGRPAKMLHLELGVGGRKASGFHNVLSGEGKGFPSVHGAG 112

DB 041 HLEPEHERRKKRGRPAKMLHLELGVGGRKASGFHNVLSGEGKGFPSVHGAG 112
 QY 113 PVACVSGSGGQSLSAISPTKKEDEG-----LDPFPGSSNASSAVV V 52
 DB 121 PVACVSGSGGQSLSAISPTKKEDEG-----LDPFPGSSNASSAVV V 52
 QY 168 LHTFHSBAAASGFVAVASSTSPASSGNSPSPGGLFAAATMGAGVAAAGTNR 265
 DB 181 LHTFHSBAAASGFVAVASSTSPASSGNSPSPGGLFAAATMGAGVAAAGTNR 265
 QY 227 SEGVAFLEGEPEFPELSEIWEWE 445
 DB 241 SEGVAFLEGEPEFPELSEIWEWE 445
 QY 287 ALKASTHEMLQVAKRNMHTETFTLEFYSKQEFPAALQVEFINDPFEFSPAME 446
 DB 301 ALKASTHEMLQVAKRNMHTETFTLEFYSKQEFPAALQVEFINDPFEFSPAME 446
 QY 347 PLGDDAFVALLIATNPSADPWNQSPPEVALGQGVVATLSTPRKRGQLEPDM 405
 DB 361 PLGDDAFVALLIATNPSADPWNQSPPEVALGQGVVATLSTPRKRGQLEPDM 405
 QY 407 LMLVSLPILSSVSEVAFPLGSKLPPLSEIWEWE 446
 DB 421 LMLVSLPILSSVSEVAFPLGSKLPPLSEIWEWE 446
 RESULT 6
 R52980
 ID R52980 standard; Protein: 461 AA.
 XX
 AC R52980;
 XX 06-OCT-1994 (first entry)
 DE Human recombinant steroid hormone receptor NER1 protein.
 XX Steroid hormone receptor; NER1.
 KW
 OS Homo sapiens.
 PN W09407916-A.
 XX
 P1 Key Location/Qualifiers
 FT Binding-site 87..154
 XX /note= "DNA binding area"
 XX
 PD 14-APR-1994.
 XX
 PF 27-SEP-1993; 93MO-0509165.
 XX
 PK 07-OCT-1992; 9208-0958137.
 XX
 PA (MERI) MERCK & CO INC.
 XX
 P1 Podan GA, Potledge SJ, Schmidt A, Vogel RG;
 XX
 DR WPI: 1994-135511/16.
 DR N-PSDB: Q63114.
 XX
 PT New human steroid receptor NER1 and corresp. DNA used to
 PT develop assays to identify steroid hormone agonists and
 PT antagonists
 PS Claim 8; Page 31; 45pp; English.
 CC The NER1 protein may be expressed in a cell culture and used
 CC to develop assays to identify steroid hormone agonists and
 CC antagonists.
 CC
 SU Sequence 461 AA;

[illegible]

RESULT 10
W25034
ID W25034 standard; protein; 460 AA.
XY

AC	W25034;	
XX		
DT	15-0001-1997	(first entry)
XX		

XX	Human ubiquitous nuclear receptor polypeptide.
KM	Ubiquitous nuclear receptor polypeptide: DNA binding domain.
KM	Therapy: drug design; diagnosis.

105 Homo sapiens.

PN US5649616-A.

PD 17-JUN-1997.

PH⁺ 10-NOV-1993; 9305-0152003.

PR 18-NOV-1994; 9405-0242411

XX

XX

XX

N-PSIHK; W25034.

PT DNA encoding ubiquitous nuclear receptor poly:polyptide(s)

XX previous and recombinant poly-peptide(s) and detecting RNA

XX
XX

receptor polypeptide (OR). The OR is useful in assays designed to select substances which interact with the OR which may potentially be of use therapeutically. OR is also used in diagnosis and drug design.

Su Sequences 460 AA;

Query Match	87.5%	Score 2023	DB 18	Length 440
Best Local Similarity	84.6%	Pred. No. 6.20	17%	
Matches 399	Conservative 18	Mismatches 28	Indels 16	Gaps 5

```
QY      1 MSSIPISLDITVINGNGSGPQPSTSAISPTIKKEGE-----IDPP-PGSGCGSSAVIV--v 5.2
       |||:||||||:||||| || :|||:|||| | || ||||: |||
DbB     1 mstpslslldlpqngqpqqatassspfvowwvpwpqvptfpdpvtpfcdvssvteglaw 6.0
```

[illegible]

```

RFSQT 11
R96235
ID R96235 standard; protein; 364 AA.

```

AC: R96235;

06-NOV-1996 (first entry)

DE Steroid hormone receptor analogue HTN small mol. variant.

Human; foetal lung; steroid hormone; receptor; analogue protein;

identifications, anticancer drug, cancerous tissues, primer.

small molecule; protein; Km

05 Homo sapiens.

PN W09609324-A1

28-MAK-1496.

XX	21-SEP-1995	15W0-JPD1909.
PH	21-SEP-1995	15W0-JPD1909.

XX 21-SEP-1994; 94JP-0226270.
PK

AA
PA
(HISA) HISAI (1.11).

PA (CANC.) CANCER INST.

XX
F-1 Nakamura Y, Saito H.

WPI; 1996-188402/19.

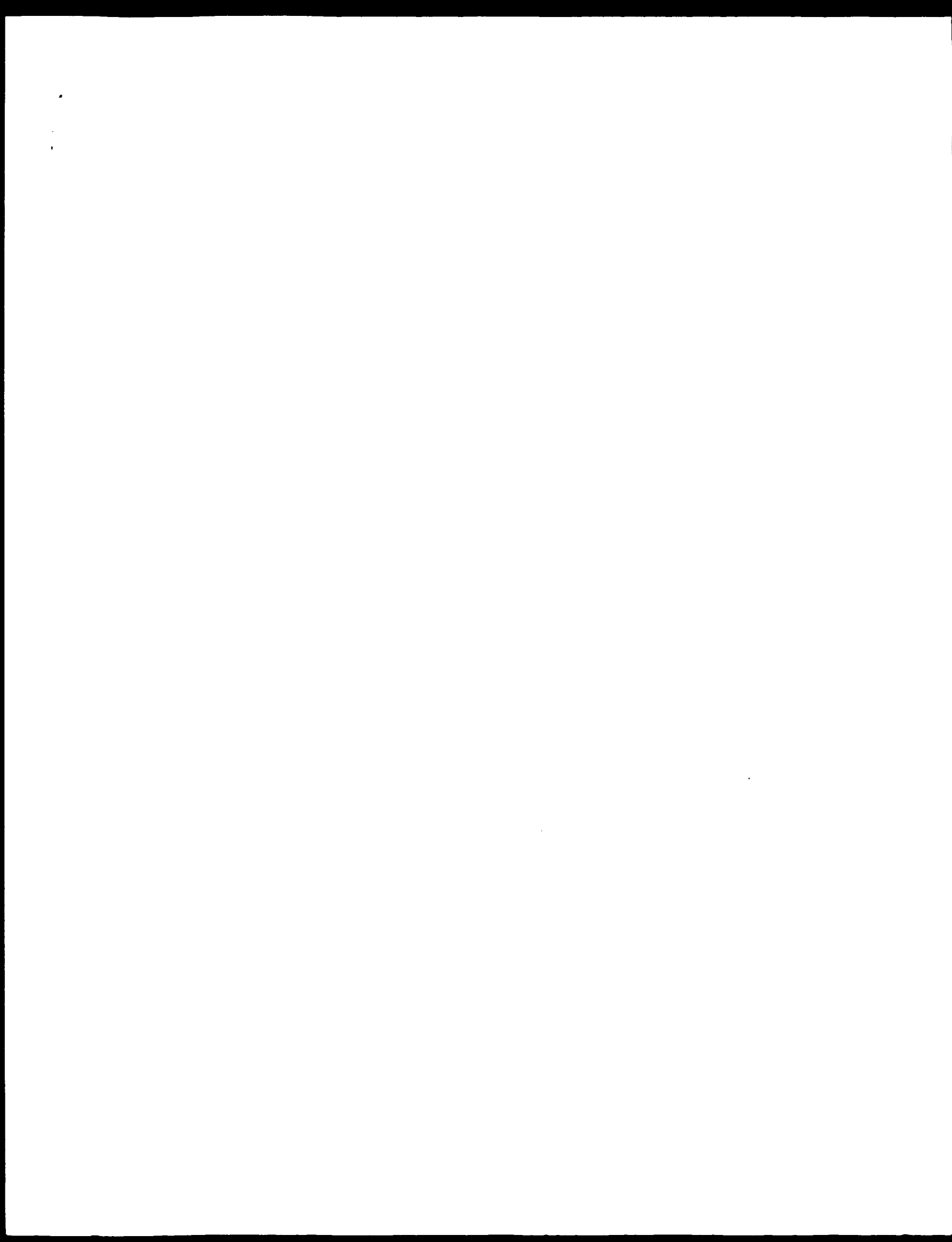
12/01/2015 15:03:15

EXON protein, and lung is expressed

XX
XX
XX

Claims 23-28; 43pp; 2013000321

90 The present sequence is the variant of the human foetal lung
91 derived steroid hormone receptor analogue protein ECDN, designated



POSTAL TELEGRAPH OFFICE OF POSTAL TELEGRAPH

TELEFAX: (713) 789-2679
 TELEX: 99-0924
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 444 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 US-08-342-411A-4

Query Match 97.24% Score 2247.51 Pos 1 Length 444
 Host Local Similarity 97.18% Pos 1 No. 740-210
 Matches 419 Conservative 6 Mismatches 4 Indels 4 Gaps 1

1 MSPTSSILTPVGNSSPQSTASPTIKKEGQTPPSSSSSAYIVLLEPEEP 60
 1 MSPTSSILTPVGNSSPQSTASPTIKKEGQTPPSSSSSAYIVLLEPEEP 57
 61 EEPKRLVAPFRLLELLELVVGLAVSLHYVNLSTGKALPSSVVDGACVAVTSSG 120
 58 EEPKRLVAPFRLLELLELVVGLAVSLHYVNLSTGKALPSSVVDGACVAVTSSG 117
 101 TLMVMTMSTVCTCTFFPTLMSTFVCTFTLFFSLCTSSGSGSLTSSASGSS 120
 116 TLMVMTMSTVCTCTFFPTLMSTFVCTFTLFFSLCTSSGSGSLTSSASGSS 117
 181 GRPASPTSTSSSSGSGPTPTGTATAGELMTGSLVAGLQCKPKSTGQKKTWHTG 240
 178 GRPASPTSTSSSSGSGPTPTGTATAGELMTGSLVAGLQCKPKSTGQKKTWHTG 237
 241 APOGSDAPGQPAHPTLALISVQPTDPAPKQVGPGLDIPPODIALKSLPTMTLQ 300
 238 APOGSDAPGQPAHPTLALISVQPTDPAPKQVGPGLDIPPODIALKSLPTMTLQ 297
 401 TARKNHETCTTKDEPTYSKDPHRAGLQVEFNPITFESKAPRGLGIDAFYALLIA 360
 298 TARKNHETCTTKDEPTYSKDPHRAGLQVEFNPITFESKAPRGLGIDAFYALLIA 357
 461 INFSAKRNVPFSSVATLQVYFALLSTPIFPPVQGLPPIPMKLKLVSLKTLSSVH 420
 358 INFSAKRNVPFSSVATLQVYFALLSTPIFPPVQGLPPIPMKLKLVSLKTLSSVH 417
 421 SEQVAFALQKLPPLSLSEIWHVE 446
 418 SEQVAFALQKLPPLSLSEIWHVE 443

RESULT 4
 US-08-330-518-2
 Sequence 2: Applicant: M. Katzarine
 Patent No. 5607967
 GENERAL INFORMATION:
 APPLICANT: Friedman, Ethan
 APPLICANT: Holloway, M. Katzarine
 APPLICANT: Rodan, Gideon
 APPLICANT: Schmidt, Azriel
 APPLICANT: Vogel, Robert
 TITLE OF INVENTION: USE OF RECEPTOR POTENTIATORS
 NUMBER OF SEQUENCES: 5
 CORRESPONDENCE ADDRESS:
 ADDRESS: Muck & Co., Inc.
 STREET: 126 East Lincoln Avenue
 CITY: Rahway
 STATE: New Jersey
 COUNTRY: US
 ZIP: 07065-0907
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US-08-330-518

CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Dolan, Catherine A.
 REGISTRATION NUMBER: 36,502
 REFERENCE/WORKSHEET NUMBER: 19416
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (908) 594-4284
 TELEFAX: (908) 594-4720
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 444 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: polypeptide
 HYPOTHEICAL: NO
 ANTI-SERIES: NO
 US-08-330-518-2

Query Match 88.04% Score 2033.51 Pos 1 Length 461
 Host Local Similarity 87.24% Pos 1 No. 476-189
 Matches 402 Conservative 17 Mismatches 27 Indels 15 Gaps 5

1 MSPTSSILTPVGNSSPQSTASPTIKKEGQTPPSSSSSAYIVLLEPEEP 51
 1 MSPTSSILTPVGNSSPQSTASPTIKKEGQTPPSSSSSAYIVLLEPEEP 60
 52 VLEPDEPE 111
 61 VLEPDEPE 120
 112 GRVAPGSGSTVMEAPDEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPE 167
 121 GRVAPGSGSTVMEAPDEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPE 160
 168 GRVAPGSGSTVMEAPDEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPE 225
 181 GRVAPGSGSTVMEAPDEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPE 240
 226 GRVAPGSGSTVMEAPDEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPE 285
 241 GRVAPGSGSTVMEAPDEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPE 300
 286 IALFASLLEIMELGIAKPNHRTSTPLFSLTYSKDPHRAGLQVEFNPITFESKAPR 345
 301 IALFASLLEIMELGIAKPNHRTSTPLFSLTYSKDPHRAGLQVEFNPITFESKAPR 340
 361 REGIDAFYALLIAINFSAKRNVPFSSVATLQVYFALLSTPIFPPVQGLPPIPMKL 420
 406 MMLKLVSLKTLSSVHSEQVAFALQKLPPLSLSEIWHVE 446
 421 MMLKLVSLKTLSSVHSEQVAFALQKLPPLSLSEIWHVE 461

RESULT 5
 US-08-330-283-2
 Sequence 2: Applicant: M. Katzarine
 Patent No. 5679518
 GENERAL INFORMATION:
 APPLICANT: Friedman, Ethan
 APPLICANT: Holloway, M. Katzarine
 APPLICANT: Rodan, Gideon
 APPLICANT: Schmidt, Azriel
 APPLICANT: Vogel, Robert
 TITLE OF INVENTION: MITOCHONDRIAL RECEPTOR POTENTIATORS
 NUMBER OF SEQUENCES: 5
 CORRESPONDENCE ADDRESS:
 ADDRESS: Muck & Co., Inc.

Qy	Db	Qy	Db
346	301	406	421
KGIGGDAEYALTAINFSA	IALTKASTIFIMLEFAPRPNHTECITFLKDETY	MLMKIVSEPTLSVSHSEGVFALRLQDEKRLPPLLS	MLMKIVSLRTLSVSHSEGVFALRLQDEKRLPPLLS
ENPWPVSPSPVA	SKDDEBPAQLQVEFIMPEFSFAM	ETLSEIWDHNE	SETIMDHNE
445	460	446	461
SYTAKKQNGNBPBP			
361			
PRGIGTAAYATFANINFSAPNPVSPPEVFA			
420			
ISYTEISPEYVAFEPFP			

RESULT 7
 PCT-US95-13924-2
 SEQUENCE 2: Amino Acid: PC/US95-13924
 GENERAL INFORMATION:
 APPLICANT: Friedman, Eitan
 APPLICANT: Holloway, M. Katherine
 APPLICANT: Kottig, Glenn
 APPLICANT: Rutledge, Su Jane
 APPLICANT: Schmidt, Arrie
 APPLICANT: Vogel, Robert
 TITLE OF INVENTION: METHOD FOR FINDING RECEPTOR POTENTIATORS
 NUMBER OF SEQUENCES: 5
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Merck & Co., Inc.
 STREET: 126 East Lincoln Avenue
 CITY: Rahway
 STATE: New Jersey
 COUNTRY: US
 ZIP: 07065-0907
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible;
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Pascal in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/13924
 FILING DATE:
 CLASSIFICATION:
 ATORNEY/AGENT INFORMATION:
 NAME: Quagliaro, Carol S.
 REGISTRATION NUMBER: 35,330
 REFERENCE/DOC. NO. NUMBER: 19427 PCT
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (908) 594 3809
 TELEFAX: (908) 594-4720
 INFORMATION FOR SEQ. ID NO. 2
 SEQUENCE CHARACTERISTICS:
 LENGTH: 461 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 HYPOTHEICAL: NO
 ANTI-SENSE: NO

[illegible][illegible]

RESULT 8
 PCT-US95-13931-2
 : Sequence 2, Application PC/TUS9513931
 : GENERAL INFORMATION:
 : APPLICANT: Friedman, Ethan
 : APPLICANT: Holloway, M. Katharine
 : APPLICANT: Rodan, Gideon
 : APPLICANT: Schmidt, Azriel
 : APPLICANT: Vargel, Robert
 : TITLE OF INVENTION: USE OF RECEPTOR POTENTIATORS
 : NUMBER OF SEQUENCES: 5
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Merck & Co., Inc.
 : STREET: 126 East Lincoln Avenue
 : CITY: Rahway
 : STATE: New Jersey
 : COUNTRY: US
 : ZIP: 07065-0607
 : MEDIUM TYPE: Floppy disk
 : COMPUTER READABLE FORM:
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent In Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER PCT/US95/13941
 : FILING DATE:
 : CLASSIFICATION:
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Quagliato, Carol S.
 : REGISTRATION NUMBER: 35,3340
 : REFERENCE/POCKET NUMBER: 19316 PCT
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (908) 594-3809
 : TELEFAX: (908) 594-4720
 : INFORMATION FOR SEQ ID NO: 2:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 461 amino acids
 : TYPE: amino acid
 : SEPARATENESS: single
 : TOPOLOGY: linear
 : MOLECULE TYPE: peptide
 : HYPOTHEICAL: NO
 : ANTI-SENSE: NO
 : PCT-US95-13931-2

Exact Match	88.0%	Score 2033.5	DB 5,	Length 461;
Best Local Similarity	87.2%	Prod. No. 4.50-189;		
Matthew's Acc.	77.0%	Similarity	271	Trials 17, Caps 5

1 INFORMATION FOR SEQ ID NO: 1:
 2 SEQUENCE CHARACTERISTICS:
 3 LENGTH: 447 amino acids
 4 TYPE: amino acid
 5 TOPOLOGY: linear
 6 MOLECULE TYPE: protein
 7 US-08-374-935-1

Query Match 60.4% Score 1346.57 DB 1 Length 447
 Host Local Similarity 62.5% Pred. No. 146-127
 Matches 283 Conservative 53 Mismatches 88 Indels 29 Gaps 6

QY 4 PPSLTPTVPPG MSPGQSTATSTPTTPPGPTTPPGS-----PSSSAVIVLLEPE 57
 DB 14 PPSAVLMPKPAQASSQAQSSQSLLEFARMPSAQTAMVHLEAPLALLTRAPP 73
 QY 7P PPP---EEFEKFNADPLJHETCFVCCAKSTFHYVNISSCKETREKRVHICANR 119
 DB 74 SEPTTETGPPRRTRAPPIHETSVGTFVASTPVAWVGGGKRISSVTKG-AM 121
 QY 114 VAKSSSTQMLAMFMRKQITREKTFACMPQVYSSSTQIKKKIKQGGQPPPPY 173
 DB 132 YCHHCHTET 192
 QY 174 PPAASSSPASPSPSPSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 233
 DB 192 LPPKSS-PPQILP-----QLSPQILMIKILVAAQCCNRRSPSLRR 234
 QY 234 VTPMDANPGRARQGFARHETLATSVAFLVMAKAVSGELQTPGQALTEAST 293
 DB 235 VTPMDANPGRARQGFARHETLATSVAFLVMAKAVSGELQTPGQALTEAST 294
 QY 294 TTHLLQAVSTNHTTETETETETETETETETETETETETETETETETETETET 353
 DB 296 TTHLLQAVSTNHTTETETETETETETETETETETETETETETETETETETET 354
 QY 354 PVALIINISAPNPVCPSPVAFQPPVAFI SVTPKTPPPVQ PTPPMETVSL 413
 DB 355 PVALIINISAPNPVCPSPVAFQPPVAFI SVTPKTPPPVQ PTPPMETVSL 414
 QY 414 KTSVHSQVAFALQPKLPILLSIMDVHE 446
 DB 415 KTSVHSQVAFALQPKLPILLSIMDVHE 447

RESULT 11
 US-08-343-358-8
 Sequence 8, Application US/08032200
 Patent No. 5571696
 GENERAL INFORMATION:
 APPLICANT: EVANS Ph.D., RONALD M.
 APPLICANT: MANGELSDORE Ph.D., DAVID J.
 APPLICANT: ONG MS., ESTELITA S.
 APPLICANT: ONG Ph.D., ANTHONY E.
 APPLICANT: BOCHMEYER Ph.D., DWE K.
 APPLICANT: GIOHRE Ph.D., VINCENT MMN
 APPLICANT: YAO M., ISSO FANG MMN
 TITLE OF INVENTION: NOVEL RECEPTORS
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESS: Preilly, Schroeder, Bergermann & Clark
 STREET: 444 So. Flower St., Suite 2000
 CITY: Los Angeles
 STATE: CA
 COUNTRY: US
 ZIP: 90071-2721
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentLib Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:

1 APPLICATION NUMBER: 08/032761, 08-
 2 FILING DATE:
 3 CLASSIFICATION: 435
 4 PRIOR APPLICATION DATA:
 5 APPLICATION NUMBER: 08/032761, 08-
 6 FILING DATE:
 7 ATTORNEY/AGENT INFORMATION:
 8 NAME: Keller Ph.D., Stephen E.
 9 REGISTRATION NUMBER: 31192
 10 REFERENCE/PACKET NUMBER: 141 8946
 11 TELECOMMUNICATION INFORMATION:
 12 TELEPHONE: (619) 535-9001
 13 TELEFAX: (619) 535-8949
 14 INFORMATION FOR SEQ ID NO: 8:
 15 SEQUENCE CHARACTERISTICS:
 16 LENGTH: 440 amino acids
 17 TYPE: amino acid
 18 TOPOLOGY: linear
 19 MOLECULE TYPE: protein
 20 US-08-252-758-8

Query Match 60.3% Score 1346 DB 1 Length 440
 Host Local Similarity 62.1% Pred. No. 76-127
 Matches 291 Conservative 54 Mismatches 83 Indels 29 Gaps 7

QY 4 PPSLTPTVPPG MSPGQSTATSTPTTPPGPTTPPGS-----PSSSAVIVLLEPE 60
 DB 14 PPSAVLMPKPAQASSQAQSSQSLLEFARMPSAQTAMVHLEAPLALLTRAPP 69
 QY 61 ---SPKPFETAPPM/INLP/CPVCKASCTHYVNISSCKETREKRVHICANR 116
 DB 70 TETPQKRAKPAKQWQNTGTFVASTPVAWVGGGKRISSVTKG-ARTIC 127
 QY 117 KQSGQDMKAPRRKQGLPKLRKFAVMPQVYSSSTQIKKKIKQGGQPPPPY 176
 DB 128 KQSGQDMKAPRRKQGLPKLRKFAVMPQVYSSSTQIKKKIKQGGQPPPPY 187
 QY 177 ASSSPRASPSPSPSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 235
 DB 188 KRSS-PPQILP-----QLSPQILMIKILVAAQCCNRRSPSLRR 239
 QY 237 VTPMDANPGRARQGFARHETLATSVAFLVMAKAVSGELQTPGQALTEAST 296
 DB 238 VTPMDANPGRARQGFARHETLATSVAFLVMAKAVSGELQTPGQALTEAST 299
 QY 297 TTHLLQAVSTNHTTETETETETETETETETETETETETETETETETETETET 356
 DB 299 TTHLLQAVSTNHTTETETETETETETETETETETETETETETETETETETET 360
 QY 357 PVALIINISAPNPVCPSPVAFQPPVAFI SVTPKTPPPVQ PTPPMETVSL 416
 DB 358 PVALIINISAPNPVCPSPVAFQPPVAFI SVTPKTPPPVQ PTPPMETVSL 417
 QY 417 KTSVHSQVAFALQPKLPILLSIMDVHE 446
 DB 418 KTSVHSQVAFALQPKLPILLSIMDVHE 449

RESULT 12
 US-08-463-694-8
 Sequence 8, Application US/08463694
 Patent No. 5696233
 GENERAL INFORMATION:
 APPLICANT: EVANS Ph.D., RONALD M.
 APPLICANT: MANGELSDORE Ph.D., DAVID J.
 APPLICANT: ONG MS., ESTELITA S.
 APPLICANT: ONG Ph.D., ANTHONY E.
 APPLICANT: BOCHMEYER Ph.D., DWE K.
 APPLICANT: GIOHRE Ph.D., VINCENT MMN
 APPLICANT: YAO M., ISSO FANG MMN
 TITLE OF INVENTION: NOVEL RECEPTORS
 NUMBER OF SEQUENCES: 14

```

CORRESPONDENCE ADDRESS:
ADDRESSER: pretty, Schroeder, Brueggemann & Clark
STREET: 444 So. Flower St., Suite 2000
CITY: Los Angeles
STATE: CA
ZIP: 90071-2921
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM pc compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patient In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/464,694
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US-07/761,064
FILING DATE: 17 SEP-1991
ATTORNEY/AGENT INFORMATION:
NAME: Keller Ph.D., Stephen E.
REGISTRATION NUMBER: 31192
REFERENCE/Docket NUMBER: P31 8943
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 545-9001
TELEFAX: (619) 545-8949
INFORMATION FOR SEQ. ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 440 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-463-694-8

Query Match      60.38: Score 1395; DB 1; Length 440;
Best Local Similarity 63.18; Prod. No. 3.7e-127;
Matches 284; Conservative 54; Mismatches 84; Indels 40; Gaps 7;

QY 4 PTSSLDTPVQ-NSPOPSISAPITIKRGOFTDPPGSSSS--AVYVILDEPDP 60
DB 14 PDSAVEIMKPKAOJASSQAQGSNCILREFAR----MPSASATAPPTALTITRAEPSPSE 69
QY 61 -----ERRKRRKPAVKMGLHETLQVNRKASPIYVNLSCDCKGFFRSVVGAGAPVAG 115
DB 70 TELPQKRRKPAVKMGLHETLQVNRKASPIYVNLSCDCKGFFRSVVGAGAPVAG 127
QY 117 KQSGTVMAMMRKPLTETLQVNRKASPIYVNLSCDCKGFFRSVVGAGAPVAG 176
DB 128 HSGGHCMDIYMRKPLTETLQVNRKASPIYVNLSCDCKGFFRSVVGAGAPVAG 187
QY 177 ASSSGPRAASPTSEASSGSGSDETEHTLQVNRKASPIYVNLSCDCKGFFRSVVGAGAPVAG 236
DB 188 KRSS-PTQITP-----QLSPGQIMTKRIKIVAAQDQGNRSPSPDRLKVP 230
QY 247 WPLCADPQSRKARQGRFAHFTLAIISVGEIVDFAKGVGEGTQITREDDQITAKSTIEI 296
DB 241 WPMADPQSRKARQGRFAHFTLAIISVGEIVDFAKGVGEGTQITREDDQITAKSTIEI 290
QY 297 MLTQIARVYNIETETLQVNRKASPIYVNLSCDCKGFFRSVVGAGAPVAG 356
DB 291 MLTQIARVYNIETETLQVNRKASPIYVNLSCDCKGFFRSVVGAGAPVAG 350
QY 457 LIAINIPSAIDQNVQPSRVEAFIQIYVVALSYITIKQVQITLQVPMIMKIVSLKIL 416
DB 451 LIAINIPSAIDQNVQPSRVEAFIQIYVVALSYITIKQVQITLQVPMIMKIVSLKIL 410
QY 417 SSVSRKQVAFIRIQKPLPLSLKIVGHE 446
DB 411 SSVSRKQVAFIRIQKPLPLSLKIVGHE 440

RESULT 14

```

```

US-08-694-501-8
Sequence 8; Application US/08694501
Patent No. 5710004
GENERAL INFORMATION:
APPLICANT: EVANS Ph.D., RONALD M.
APPLICANT: MANGELSDORF Ph.D., DAVID J.
APPLICANT: ONG MS., ESTELITA S.
APPLICANT: ONG Ph.D., ANTHONY E.
APPLICANT: HOSCHKEVER Ph.D., TWE K.
APPLICANT: GICHIRE Ph.D., VINCENT NNN
APPLICANT: YAO M., TSO-PAN
TITLE OF INVENTION: NOVEL REPERTOIRS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSER: pretty, Schroeder, Brueggemann & Clark
STREET: 444 So. Flower St., Suite 2000
CITY: Los Angeles
STATE: CA
ZIP: 90071-2921
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM pc compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patient In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/464,694
FILING DATE: 07-AUG-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/444,358
FILING DATE:
APPLICATION NUMBER: US/07/761,068
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Keller Ph.D., Stephen E.
REGISTRATION NUMBER: 31192
REFERENCE/Docket NUMBER: P31 8946
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ. ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 440 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-694-501-8

Query Match      60.38: Score 1395; DB 1; Length 440;
Best Local Similarity 63.18; Prod. No. 3.7e-127;
Matches 284; Conservative 54; Mismatches 84; Indels 40; Gaps 7;

QY 4 PTSSLDTPVQ-NSPOPSISAPITIKRGOFTDPPGSSSS--AVYVILDEPDP 60
DB 14 PDSAVEIMKPKAOJASSQAQGSNCILREFAR----MPSASATAPPTALTITRAEPSPSE 69
QY 61 -----ERRKRRKPAVKMGLHETLQVNRKASPIYVNLSCDCKGFFRSVVGAGAPVAG 115
DB 70 TELPQKRRKPAVKMGLHETLQVNRKASPIYVNLSCDCKGFFRSVVGAGAPVAG 127
QY 117 KQSGTVMAMMRKPLTETLQVNRKASPIYVNLSCDCKGFFRSVVGAGAPVAG 176
DB 128 HSGGHCMDIYMRKPLTETLQVNRKASPIYVNLSCDCKGFFRSVVGAGAPVAG 187
QY 177 ASSSGPRAASPTSEASSGSGSDETEHTLQVNRKASPIYVNLSCDCKGFFRSVVGAGAPVAG 236
DB 188 KRSS-PTQITP-----QLSPGQIMTKRIKIVAAQDQGNRSPSPDRLKVP 230
QY 247 WPLCADPQSRKARQGRFAHFTLAIISVGEIVDFAKGVGEGTQITREDDQITAKSTIEI 296
DB 241 WPMADPQSRKARQGRFAHFTLAIISVGEIVDFAKGVGEGTQITREDDQITAKSTIEI 290

```


QY 207 M1QIAPPPNHTPOTITKPTYSKIDPBPACI QVEEINIPERSAMPPLIETONAYE 256
 DB 291 M1ETSPNPSFSTITLKTENEDFARAGLVERINIPFSPAMNLOINDAEFA 350
 QY 357 TITAINPSPAPNPVPSPPVPAI QDPVEALLSTPTEKPPQDOLPPPMKIVSLPTL 416
 DB 451 TITAINPSPAPNPVQDQI QVRIQHTYVEAI HAVSTIHHRIDLMFPRMLKLSLPTL 410
 QY 417 SSVSPQVPAIPI QDKKLPLLSLMDVHE 446
 DB 411 SSVSPQVPAIPI QDKKLPLLSLMDVHE 440

RESULT 14
 US-08-466-120-2
 Sequence 2: Application No. US/08466120
 Patent No. 5869284
 GENERAL INFORMATION:
 APPLICANT: CAO, ET AL.
 TITLE OF INVENTION: Peptideic Acid Receptor Epsilon
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: CARELLA, RYNE, BAIN, GILFILLAN,
 STREET: 6 BECKER FARM ROAD
 CITY: ROSHLAND
 STATE: NEW JERSEY
 COUNTRY: USA
 ZIP: 07068
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5 INCH DISKETTE
 COMPUTER: IBM PS/2
 OPERATING SYSTEM: MS-DOS
 SOFTWARE: WORD PERFECT 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08466120
 FILING DATE: June 6, 1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US94/07266
 FILING DATE: 24 JUN 94
 ATTORNEY/AGENT INFORMATION:
 NAME: FERRARO, GREGORY D.
 REGISTRATION NUMBER: 36,134
 REFERENCE/AGENT NUMBER: 325600-354
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 201-994-1700
 TELEFAX: 201-994-1744
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 433 AMINO ACIDS
 TYPE: AMINO ACID
 STRANDEDNESS:
 TOPOLOGY: LINEAR
 MOLECULE TYPE: PROTEIN
 US-08-466-120-2

Query Match 55.4% Score 1281.5 DB 2 Length 433
 Best Local Similarity 58.7% Pred. No. 3,76-116
 Matches 266 Conservative 52 Mismatches 92 Indels 43 Gaps 7

QY 4 PTSGSLTPVPG NSTQNTSAISLTKFPGSTTPPGS FSSSSAYIVILEPE 57
 DB 14 POSAVELMKPGAGMGAGGSCSTIFEFAMPHSAGIAGVLEAEPTALLTPAEP 73
 QY 50 PPEP-----PPKQAPYPLGHPVPGVQVASKCHYVWLSQEGKPPESVWGACNE 112
 DB 74 SETELPRPKPKKPAVPMKINSISVQVQKASCFHYVWLSQEGKPPESVWIKQ-AN 131
 QY 114 YATPSTSTQMTAKESPKTQICGCFKFAVMPFQVLSFSLKPPKQVQVQVQVQV 172

DB 132 YTHSSGRTWKTVMPPKQ-----FVVI SEFGITPKETKQPPQVHATS 177
 QY 174 PPAASSQPPAASPPPTSPASSQSGSSTSTTPT TAQPTMQQVAAQT GQNKSSSQPK 233
 DB 178 LPPRASS-PPQITP-----GIPQGLAMIFKIVPAQDQGNKSSFSRLR 220
 QY 234 VTHNIDAMVGGELAKGGVPAHIFLAIVYFVGAFAVNTETLQCTETGVALFAST 263
 DB 221 VTHNIDAMVGGELAKGGVPAHIFLAIVYFVGAFAVNTETLQCTETGVALFAST 260
 QY 294 TETMLQVAPNHTPOTITKPTYSKIDPBPACI QVEEINIPERSAMPPLIETONAYE 353
 DB 281 TETMLQVAPNHTPOTITKPTYSKIDPBPACI QVEEINIPERSAMPPLIETONAYE 340
 QY 354 TITAINPSPAPNPVPSPPVPAI QDPVEALLSTPTEKPPQDOLPPPMKIVSLPTL 413
 DB 341 TITAINPSPAPNPVQDQI QVRIQHTYVEAI HAVSTIHHRIDLMFPRMLKLSLPTL 400
 QY 414 SSVSPQVPAIPI QDKKLPLLSLMDVHE 446
 DB 401 SSVSPQVPAIPI QDKKLPLLSLMDVHE 433

RESULT 15
 PCT-US94-07266-2
 Sequence 2: Application No. PCT/US9407266
 GENERAL INFORMATION:
 APPLICANT: CAO, ET AL.
 TITLE OF INVENTION: Peptideic Acid Receptor Epsilon
 NUMBER OF SEQUENCES: 2
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: CARELLA, RYNE, BAIN, GILFILLAN,
 STREET: 6 BECKER FARM ROAD
 CITY: ROSHLAND
 STATE: NEW JERSEY
 COUNTRY: USA
 ZIP: 07068
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5 INCH DISKETTE
 COMPUTER: IBM PS/2
 OPERATING SYSTEM: MS-DOS
 SOFTWARE: WORD PERFECT 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US94/07266
 FILING DATE: Concurrently
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: FERRARO, GREGORY D.
 REGISTRATION NUMBER: 36,134
 REFERENCE/AGENT NUMBER: 325600-125
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 201-994-1700
 TELEFAX: 201-994-1744
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 433 AMINO ACIDS
 TYPE: AMINO ACID
 STRANDEDNESS:
 TOPOLOGY: LINEAR
 MOLECULE TYPE: PROTEIN
 PCT-US94-07266-2

Query Match 55.4% Score 1281.5 DB 2 Length 433
 Best Local Similarity 58.7% Pred. No. 3,76-116
 Matches 266 Conservative 52 Mismatches 92 Indels 43 Gaps 7

QY 4 PTSGSLTPVPG NSTQNTSAISLTKFPGSTTPPGS-----FSSSSAYIVILEPE 57
 DB 14 POSAVELMKPGAGMGAGGSCSTIFEFAMPHSAGIAGVLEAEPTALLTPAEP 73

DB 14 PUSAVEIMKPPUJAJUSJAGGSSST11PEFAPMPSAGTACVAFPAFFTALITPAEPP 73
 QY 58 DEPPER-----KKKKIATKMIIGHEUJFWGJIKASJAHNVLSJGQGGFFRPSVHGAGH 113
 DB 74 SHEETIRPPKPKKGAAPKMAENLASVAGJASGPHYVLSJGQKKFKKRSVTKG--AH 131
 QY 114 YACRGSSTFQMDAPRKKQVQJGJELKKNKAJMEHOCVLSSEJIKKKRLUKQGGQPPPS 173
 DB 132 YICHSRGHCPNNITMKKKCY-----KCVLSHEJILIKKIKKQHEEJYHATS 177
 QY 174 EDMASSSJPDAASQSTSPASSQJNPPHOCJGJYAGPJMUGJVAAGJGNNKPSSTQPK 233
 DB 178 LPPRASS-PPQTH-----QJNPEJGJMIKTI VPAQGGNNKPSSTQPLP 220
 QY 234 VPPWJAZDQSGKPAKQJPAHPTTALISVQJTVPAKQVSPHJQJSPEDJALIKAST 293
 DB 221 VTPWJMAHJHRSRKAQJQJPAHPTTALISVQJTVPAKQJPEHJQJSPEDJALIKISA 280
 QY 294 IEIMJLGTAPRYNHPTPTPTPKPTYSKJQJHJATJQVFFNPIFEESPAMPPLGLIDA 353
 DB 281 TEVMIJETSRYNIGSSTTFELKQFSYNRKHQPAKAGJQVFFINPIFEESKAMNELQJNDP 340
 QY 354 EXALLIATINISADBPVNGJPSKVPFALJQJYVAFALLSYTRIKPQJQJPEPPMIMKLVSJ 413
 DB 341 EFPALITASTISADBPVNGJQJYVFPJQJHYVFPALIDAVVSIMHHPHPTMEPPMIMKLVSJ 400
 QY 414 KJTSVHSJQVFAJRLQJRKJLPJLSEIMQVHE 446
 DB 401 KJTSVHSJQVFAJRLQJRKJLPJLSEIMQVHE 433

Search completed: April 25, 2001, 09:56:16
 Job time: 48 sec

GenCore version 4.5
Copyright (c) 1994 - 2000 CompuGen Ltd.

0M protein - protein search, using SW model

Run on: April 25, 2001, 09:55:28 / Score: 1.16, 18 Seconds
(without alignment)

1598 041 Millions cell updates/sec

Title: US-09-365-576-3

Percent Score: 29.12

Sequence: 1 MSPTSSITITVPMNSMSP

1274PKTPTTSPHWYV 446

Scoring table:

BLAST0M62
Gapop 10.0, expect 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: 1: PIR:67:*

2: PIR:1:*

3: PIR:2:*

4: PIR:4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2309	99.9	446	149021	retinol X receptor
2	2294	98.8	446	150264	retinol X receptor
3	2033.5	88.0	447	104014	steroid hormone
4	1395.5	60.4	447	148975	nuclear orphan receptor
5	1379.5	59.7	445	156043	steroid hormone receptor
6	708.5	40.6	878	141055	retinol X receptor
7	704	40.4	536	156590	retinol X receptor
8	638.5	37.4	444	146494	retinol X receptor
9	633.5	37.4	469	156918	retinol X receptor
10	582	35.2	425	157336	retinol X receptor
11	575.5	34.9	429	157229	retinol X receptor
12	570.5	34.8	429	141977	retinol X receptor
13	571	34.7	462	142943	retinol X receptor
14	563	34.4	462	156920	retinol X receptor
15	557	34.1	459	156920	retinol X receptor
16	556.5	34.1	454	143363	retinol X receptor
17	556.5	34.1	458	144714	retinol X receptor
18	555.5	34.0	555	144045	retinol X receptor
19	555	34.0	586	144197	retinol X receptor
20	551.5	33.9	443	145991	retinol X receptor
21	549	33.7	464	156558	retinol X receptor
22	548	33.7	467	156558	retinol X receptor
23	546	33.4	423	141761	retinol X receptor
24	546	33.4	423	141761	retinol X receptor
25	545.5	33.4	448	156451	retinol X receptor
26	545	33.3	452	156451	retinol X receptor
27	543.5	33.2	456	156451	retinol X receptor
28	542	33.4	454	156451	retinol X receptor
29	540	33.4	453	156451	retinol X receptor

ALIGNMENTS

Query Match	Similarity	Score	DB ID	Length	Index	Gap
1	99.9%	2309	149021	446	0	0
2	98.8%	2294	150264	446	0	0
3	88.0%	2033.5	104014	447	0	0
4	60.4%	1395.5	148975	447	0	0
5	59.7%	1379.5	156043	445	0	0
6	40.6%	708.5	141055	878	0	0
7	40.4%	704	156590	536	0	0
8	37.4%	638.5	146494	444	0	0
9	37.4%	633.5	156918	469	0	0
10	35.2%	582	157336	425	0	0
11	34.9%	575.5	157229	429	0	0
12	34.8%	570.5	141977	429	0	0
13	34.7%	571	142943	462	0	0
14	34.4%	563	156920	462	0	0
15	34.1%	557	156920	459	0	0
16	34.1%	556.5	143363	454	0	0
17	34.1%	556.5	144714	458	0	0
18	34.0%	555.5	144045	555	0	0
19	34.0%	555	144197	586	0	0
20	33.9%	551.5	145991	443	0	0
21	33.7%	549	156558	464	0	0
22	33.7%	548	156558	467	0	0
23	33.4%	546	141761	423	0	0
24	33.4%	546	141761	423	0	0
25	33.4%	545.5	156451	448	0	0
26	33.3%	545	156451	452	0	0
27	33.2%	543.5	156451	456	0	0
28	33.4%	542	156451	454	0	0
29	33.4%	540	156451	453	0	0

Db 461 INFSARPVWVDSVEALQGVYVALLSYTRKRDQDLKRPDKMLKIVSLRILSSVH 420
 QY 421 SPVVAIRKQKKIPPIISETIMVHP 446
 Db 421 SPVVAIRKQKKIPPIISETIMVHP 446

RESULT 2
 159454
 orphan nuclear receptor OK 1 - rat

C:Species: Rattus norvegicus (Norway rat)
 C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1999
 C:Accession: 159454; 159264
 R:Johnson, M.; Emswiler, P.; Li, Y.; Wikstrom, A.; Petro-Huikko, M.; Gustafsson, J. A.
 Proc. Natl. Acad. Sci. U.S.A. 92: 2089-2093, 1995
 A:Title: OK-1, a member of the nuclear receptor superfamily that interacts with the 5'- α
 A:Reference number: 159454, MIMD:95199238
 A:Accession: 159454
 A:Status: Translated from OR/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-446 <RES>

A:Expectimtal source: Spontaneous binding, hepatic
 R:Song, C.; Kokontis, J.M.; Mupparak, R.A.; Rao, S.
 Proc. Natl. Acad. Sci. U.S.A. 91: 10809-10813, 1994
 A:Title: Ubiquitous receptor: a receptor that modulates gene activation by retinoic acid
 A:Reference number: 159264; MIMD:95062154
 A:Accession: 159264
 A:Status: Translated from GB/EMBL/DBJ

A:Molecule type: mRNA
 A:Residues: 1-32; V, 34-51, 55-218, V, 220-446 <RES>
 A:Cross-references: EMBL:014543; NID:9556617; PDB:AAA52361.1; PDB:955662
 A:Expectimtal source: void
 A:Family: unassigned orba related proteins: orba transforming protein homology
 C:Keywords: DNA binding; zinc finger
 F:76-466/transformin orba transforming protein homology - orba

Query Match 98.8%; Score 2284; DB 2; Length 446;
 Best Local Similarity 98.2%; Pred. No. 4, 60-156;
 Matches 400; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSSPSSIPPPVGNNSPPSSATSPTRKPGQEHDPDPSGSSAVIVVLEPEDP 60
 Db 1 MSSPSSIPPPVGNNSPPSSATSPTRKPGQEHDPDPSGSSAVIVVLEPEDP 60
 QY 61 EKKKKRPKKMLQHPGCVTKKASPIYNYVSTPGCGFPSSVVRGZAPVAPPSG 120
 Db 61 EKKKKRPKKMLQHPGCVTKKASPIYNYVSTPGCGFPSSVVRGZAPVAPPSG 120
 QY 121 TQDMALPRKKQGLVYKPKKFAEMREGVLSPEQIPKKRTPKQGGQPPSPSEPAASS 180
 Db 121 TQDMALPRKKQGLVYKPKKFAEMREGVLSPEQIPKKRTPKQGGQPPSPSEPAASS 180
 QY 181 GFAASSTSPASSGSGSGEGGQTLTADELMLQGLVAQLQNKSTFSDQKVPWPLG 240
 Db 181 GFAASSTSPASSGSGSGEGGQTLTADELMLQGLVAQLQNKSTFSDQKVPWPLG 240
 QY 241 ADPQSPAPGQPPPAHPTFAISVAPTVPAKQVPSFELGSEPDQTALEKASTELMELQ 300
 Db 241 ADPQSPAPGQPPPAHPTFAISVAPTVPAKQVPSFELGSEPDQTALEKASTELMELQ 300
 QY 401 FARKYNIETETTELKQFYSKDQFRAQLVFEINPFEESPARMLGLDAEFALLIA 360
 Db 401 FARKYNIETETTELKQFYSKDQFRAQLVFEINPFEESPARMLGLDAEFALLIA 360
 QY 461 INFSARPVWVDSVEALQGVYVALLSYTRKRDQDLKRPDKMLKIVSLRILSSVH 420
 Db 461 INFSARPVWVDSVEALQGVYVALLSYTRKRDQDLKRPDKMLKIVSLRILSSVH 420
 QY 421 SPVVAIRKQKKIPPIISETIMVHP 446
 Db 421 SPVVAIRKQKKIPPIISETIMVHP 446

RESULT 3

stereoid hormone nuclear receptor NR1 - human

C:Species: Homo sapiens (man)
 C:Date: 13-Jun-1995 #sequence_revision 14-Jul-1995 #text_change 24-Sep-1999
 C:Accession: J04014

R:Shinar, D.M.; Endo, N.; Rulledge, S.J.; Vogel, K.; Rodan, G.A.; Schmidt, A.
 Gene 147, 273-276, 1994
 A:Title: NR1, a new member of the gene family encoding the human steroid hormone nucl

A:Reference number: J04014; MIMD:95011628
 A:Accession: J04014
 A:Molecule type: mRNA
 A:Residues: 1-443 <RES>
 A:Cross-references: GB:007132; NID:9641961; PDB:AAA61783.1; PDB:9641962
 A:Experimental source: osteosarcoma cells SMS Z/810
 C:Genetics:

A:Gene: GDB:DNK
 A:Cross-references: GDB:389570; GIMM:600380
 A:Map position: 19q13.3-19q13.4
 C:Expectimtal source: unassigned orba related proteins: orba transforming protein homology
 C:Keywords: steroid hormone receptor
 F:85-381/omatin: orba transforming protein homology - ERBA-
 F:87-134/omatin: DNA binding #status predicted - RN

Query Match 88.0%; Score 2043.5; DB 2; Length 401;
 Best Local Similarity 87.2%; Pred. No. 40-148;
 Matches 402; Conservative 17; Mismatches 27; Indels 15; Gaps 5;

QY 1 MSSPSSIPPPVGNNSPPSSATSPTRKPGQEHDPDPSGSSAVIVVLEPEDP 51
 Db 1 MSSPSSIPPPVGNNSPPSSATSPTRKPGQEHDPDPSGSSAVIVVLEPEDP 60
 QY 52 VVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPV 111
 Db 52 VVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPV 120
 QY 112 GRYVAGPSPGQMDAPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPV 167
 Db 112 GRYVAGPSPGQMDAPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPV 180
 QY 121 KRYVAGPSPGQMDAPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPV 180
 Db 121 KRYVAGPSPGQMDAPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPV 190
 QY 168 --GPPSPHMAASSGSPASAPSPISASSGSGSGEGGQTLTADELMLQGLVAQLQNK 225
 Db 168 --GPPSPHMAASSGSPASAPSPISASSGSGSGEGGQTLTADELMLQGLVAQLQNK 240
 QY 181 GSGSPVAGGSSSSASAPSPISASSGSGSGEGGQTLTADELMLQGLVAQLQNK 240
 Db 181 GSGSPVAGGSSSSASAPSPISASSGSGSGEGGQTLTADELMLQGLVAQLQNK 240
 QY 226 PSPVGGKVPWPLGAPVSPAPGQPPPAHPTFAISVAPTVPAKQVPSFELGSEPD 285
 Db 226 PSPVGGKVPWPLGAPVSPAPGQPPPAHPTFAISVAPTVPAKQVPSFELGSEPD 300
 QY 241 PSPVGGKVPWPLGAPVSPAPGQPPPAHPTFAISVAPTVPAKQVPSFELGSEPD 300
 Db 241 PSPVGGKVPWPLGAPVSPAPGQPPPAHPTFAISVAPTVPAKQVPSFELGSEPD 300
 QY 286 IALLKASTELMELQAPVNIETETTELKQFYSKDQFRAQLVFEINPFEESPAR 345
 Db 286 IALLKASTELMELQAPVNIETETTELKQFYSKDQFRAQLVFEINPFEESPAR 360
 QY 301 IALLKASTELMELQAPVNIETETTELKQFYSKDQFRAQLVFEINPFEESPAR 360
 Db 301 IALLKASTELMELQAPVNIETETTELKQFYSKDQFRAQLVFEINPFEESPAR 360
 QY 346 KRLGLDAEFALLIAINFSARPVWVDSVEALQGVYVALLSYTRKRDQDLKRP 405
 Db 346 KRLGLDAEFALLIAINFSARPVWVDSVEALQGVYVALLSYTRKRDQDLKRP 420
 QY 406 MIMKIVSLRILSSVSHSROVFAIRLQKKLPIILSILMIVHP 446
 Db 406 MIMKIVSLRILSSVSHSROVFAIRLQKKLPIILSILMIVHP 446
 Db 421 MIMKIVSLRILSSVSHSROVFAIRLQKKLPIILSILMIVHP 446

RESULT 4

orphan nuclear receptor LXR-alpha - human

C:Species: Homo sapiens (man)
 C:Date: 16-Feb-1996 #sequence_revision 16-Feb-1996 #text_change 20-Sep-1999
 C:Accession: L38975

R:Willi, P.; T. Desvergne, P.; Ouel, E.S.; Evans, R.M.; Heyman, R.A.; Mangelsdorf, D.J.
 Genes Dev. 9, 1043-1045, 1995
 A:Title: LXR, a nuclear receptor that defines a distinct retinoid response pathway.

265 InduileValaspheAlaValsgInValProlylPheLeuGlnLeuGly 281
 551 ACCAGATCGGAGACTGCTTAACCAAGTCCCTGCTTCTGTAACCTGGC 600
 282 ATGCLASPheLeuAlaValLeuGlnValAspThrTleGlnIleMetLe 298
 601 CGGAGAGCCAAATCGCCCTGCTGAAGGATCGACTATATACATCAT 650
 298 ATGCGTATATATATATATATATATATATATATATATATATATAT 615
 651 TTTAAGGATATATATATATATATATATATATATATATATATATAT 698
 315 GGLysAspPheThrTyrSerTyrAspAspPhe 325
 699 TTAAGGACTTCATCATACAGCAATACAGCACTTC 730
 seq_name: qb_est7:1BE02387

seq_documentation_block:
 LOCUS BE382387 721 bp mRNA EST 21-JUN-2000
 DEFINITION G312966341 NIH_MGC_19 Homo sapiens: LMA clone IMAGE:362894 5'
 mRNA sequence.
 ACCESSION BE382387
 VERSION BE382387.1 01/03/27752
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo;

REFERENCE
 1 (bases 1 to 721)
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Tel: (401) 456-1550
 Email: robert_strausberg@nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Using Homopolylic Subtraction
 cDNA Library Arrayed by: The L.M.A.G.E. Consortium (LINC)
 DNA Sequencing by: Incycle Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the L.M.A.G.E. Consortium/INL at: image.linc.nih.gov
 Plate: LINC314 row: 1 column: 23
 High quality sequence: 674.
 Location/Qualifiers
 1..721
 /organism "Homo sapiens"
 /db_xref "taxon:9606"
 /clone "IMAGE:362894"
 /clone_1 "NIH_MGC_19"
 /tissue_type "neuroblastoma"
 /lab_host "B1010R (phage-resistant)"
 /vector "pUC19 (phage-resistant)"
 /notes "cDNA made by oligo-dT priming. Effectively
 cloned into EcoRI/XbaI sites using the following 5'
 adaptor: GCGACGAGCTC. Library constructed by using Homo
 in the laboratory of Terada & Watanabe (University of
 California, Berkeley) using ZAP-cDNA system (Stratagene)
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: This is a NIH_MGC Library."

FEATURES
 source
 1..721
 /organism "Homo sapiens"
 /db_xref "taxon:9606"
 /clone "IMAGE:362894"
 /clone_1 "NIH_MGC_19"
 /tissue_type "neuroblastoma"
 /lab_host "B1010R (phage-resistant)"
 /vector "pUC19 (phage-resistant)"
 /notes "cDNA made by oligo-dT priming. Effectively
 cloned into EcoRI/XbaI sites using the following 5'
 adaptor: GCGACGAGCTC. Library constructed by using Homo
 in the laboratory of Terada & Watanabe (University of
 California, Berkeley) using ZAP-cDNA system (Stratagene)
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: This is a NIH_MGC Library."

alignment_scores:
 quality: 964.00 length: 193
 ratio: 4.995 gaps: 0
 percent similarity: 100.000 percent identity: 98.964
 alignment_block:
 us-09-365-576-2 x BE382387

Align: us-09-365-576-2 BE382387 1 EST: 1 to 721

254 AlaHisPheThrGluLeuAlaValIleSerValInduileValAspPhe 270
 4 CGCCACTTTCAGGAGCTGGCTATCATCTTAATTATAGAGATCGCGATT 52
 270 GAlaLysGlnValProlylPheLeuGlnLeuGlyAspGlnIleVal 287
 53 CGCTAAATCAAGTCCCTGCTGCTGAAGGATCGACTATATACATCAT 102
 287 AlaLeuGlnValAspThrTleGlnIleMetLeuGlnIleThrAlaG 303
 103 CGCTGCTAAAGATATATATATATATATATATATATATATATATAT 152
 304 ATGTYTAsnHisGlnThrGluSerTleThrPheGlnLysAspPheThr 320
 153 CGCTATCAATCAATACAGCACTGATATCACTTCTGAGGACTTCAGCTA 202
 320 TSeTyrAspAspPheThrSArgAlaGlySerGlnValGlnIleAsp 337
 203 CAGCAGAGCAAGCTTCACCGCTGCAAGGATATATATATATATATATAT 252
 337 TAlaPheGlnThrSerATATATATATATATATATATATATATATAT 353
 253 CATCTTCAGATTCCTGGAGAGCATATATATATATATATATATATAT 302
 354 GAlaTAT 370
 303 GATATACCTGCTGCTATATATATATATATATATATATATATATATAT 352
 370 AT 487
 353 CGCGAGAGCAAGCTTCACCGCTGCAAGGATATATATATATATATAT 402
 487 AlaLeuGlnValProlylPheLeuGlnLeuGlyAspGlnIleVal 493
 403 CGCTAAATCAAGTCCCTGCTGCTGAAGGATCGACTATATACATCAT 452
 493 ATGCLASPheLeuAlaValLeuGlnValAspThrTleGlnIleMet 420
 453 CGCGAGAGCAAGCTTCACCGCTGCAAGGATATATATATATATATAT 502
 420 SSeTyrAspAspPheThrSArgAlaGlySerGlnValGlnIleAsp 437
 503 CGCGAGAGCAAGCTTCACCGCTGCAAGGATATATATATATATATAT 552
 437 GLeuSerGlnValLeuPAspValHisGln 446
 553 TCGCTGAGAGATCGAGGATCGAGGAT 581

seq_name: qb_est7:1BE391856

seq_documentation_block:
 LOCUS BE391856 729 bp mRNA EST 21-JUN-2000
 DEFINITION G312966341 NIH_MGC_44 Homo sapiens: cDNA clone IMAGE:561514 5'
 mRNA sequence.

ACCESSION BE391856
 VERSION BE391856.1 01/03/27221
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo;

REFERENCE
 1 (bases 1 to 729)
 AUTHORPE
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Tel: (401) 456-1550
 Email: robert_strausberg@nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Using Homopolylic Subtraction
 cDNA Library Arrayed by: The L.M.A.G.E. Consortium (LINC)


```
|||||
452 AGTCTCCAGCAGCTTCCTCAGCAGCAGCTTTGGCCACTTCACTCAGCTAGCT 403
261 TTCTTCTGCTGATCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 277
|||||
402 ATGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 453
277 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 294
|||||
352 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 403
294 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 310
|||||
302 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 253
311 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 427
|||||
252 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 203
327 GATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 444
|||||
202 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 153
344 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 460
|||||
152 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 103
361 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 477
|||||
102 ATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 53
377 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 494
|||||
52 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4
394 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 494
|||||
2 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1
```

4

5

6

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 29, 2001, 14:15:27 ; Search time 141.69 Seconds
(without alignments)
204.151 Million cell updates/sec

Title: US-09-365-576-7

Perfect score: 1860
Sequence: 1 GCGAGGGGACATAGTGGGAAAAAAAAAAAAAAAAAAAA 1860

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapov 1.0

Scorepool: 9623517 score, 74081774 residues

Total number of hits satisfying chosen parameters: 19247034

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database :

EST:*

1: qb_est1: *
2: qb_est2: *
3: qb_est3: *
4: qb_est4: *
5: qb_est5: *
6: qb_est6: *
7: qb_est7: *
8: qb_est8: *
9: qb_est9: *
10: qb_est10: *
11: qb_est11: *
12: qb_est12: *
13: qb_est13: *
14: qb_est14: *
15: qb_est15: *
16: qb_est16: *
17: qb_est17: *
18: qb_est18: *
19: qb_est19: *
20: qb_est20: *
21: qb_est21: *
22: qb_est22: *
23: qb_est23: *
24: qb_est24: *
25: qb_est25: *
26: qb_est26: *
27: qb_est27: *
28: qb_est28: *
29: qb_est29: *
30: qb_est30: *
31: qb_est31: *
32: qb_est32: *
33: qb_est33: *
34: qb_est34: *
35: qb_est35: *
36: qb_est36: *
37: qb_est37: *
38: qb_est38: *
39: qb_est39: *
40: qb_est40: *
41: qb_est41: *
42: qb_est42: *
43: qb_est43: *

44: em_esthum10: *
45: em_esthum11: *
46: em_esthum12: *
47: em_esthum13: *
48: em_esthum14: *
49: em_esthum15: *
50: em_esthum16: *
51: em_esthum17: *
52: em_esthum18: *
53: em_esthum19: *
54: em_esthum20: *
55: em_esthum21: *
56: em_esthum22: *
57: em_esthum23: *
58: em_esthum24: *
59: em_esthum25: *
60: em_esthum26: *
61: em_esthum27: *
62: em_esthum28: *
63: em_est101: *
64: em_est102: *
65: em_est103: *
66: em_est104: *
67: em_est105: *
68: em_est106: *
69: em_est107: *
70: em_est108: *
71: em_est109: *
72: em_est110: *
73: em_est111: *
74: em_est112: *
75: em_est113: *
76: em_est114: *
77: em_est115: *
78: em_est116: *
79: em_est117: *
80: em_est118: *
81: em_est119: *
82: em_est120: *
83: em_est121: *
84: em_est122: *
85: em_est123: *
86: em_est124: *
87: em_est125: *
88: em_est126: *
89: em_est127: *
90: em_est128: *
91: em_est129: *
92: em_est130: *
93: em_est131: *
94: em_est132: *
95: em_est133: *
96: em_est134: *
97: em_est135: *
98: em_est136: *
99: em_est137: *
100: em_est138: *
101: em_est139: *
102: em_est140: *
103: em_est141: *
104: em_est142: *
105: em_est143: *
106: em_est144: *
107: em_est145: *
108: em_est146: *
109: em_est147: *
110: em_est148: *
111: em_est149: *
112: em_est150: *
113: em_est151: *
114: em_est152: *
115: em_est153: *
116: em_est154: *

117: qb_est48:*
 118: qb_est49:*
 119: qb_est50:*
 120: qb_est51:*
 121: qb_est52:*
 122: qb_est53:*
 123: qb_est54:*
 124: qb_est55:*
 125: qb_est56:*
 126: qb_est57:*
 127: qb_est58:*
 128: qb_est59:*
 129: qb_est60:*
 130: qb_est61:*
 131: qb_est62:*
 132: qb_est63:*
 133: qb_est64:*
 134: qb_est65:*
 135: qb_est66:*
 136: qb_est67:*
 137: qb_est68:*
 138: qb_est69:*
 139: qb_est70:*
 140: qb_est71:*
 141: qb_est72:*
 142: qb_est73:*
 143: qb_est74:*
 144: qb_est75:*
 145: qb_est76:*
 146: qb_est77:*
 147: qb_est78:*
 148: qb_est79:*
 149: qb_est80:*
 150: qb_est81:*
 151: qb_est82:*
 152: qb_est83:*
 153: qb_est84:*
 154: qb_est85:*
 155: qb_est86:*
 156: qb_est87:*
 157: qb_est88:*
 158: qb_est89:*
 159: qb_est90:*
 160: qb_est91:*
 161: qb_est92:*
 162: qb_est93:*
 163: qb_est94:*
 164: qb_est95:*
 165: qb_est96:*
 166: qb_est97:*
 167: qb_est98:*
 168: qb_est99:*
 169: qb_est100:*
 170: qb_est101:*
 171: qb_est102:*
 172: qb_est103:*
 173: qb_est104:*
 174: qb_est105:*
 175: qb_est106:*
 176: qb_est107:*
 177: qb_est108:*
 178: qb_est109:*
 179: qb_est110:*
 180: qb_est111:*
 181: qb_est112:*
 182: qb_est113:*
 183: qb_est114:*
 184: qb_est115:*
 185: qb_est116:*
 186: qb_est117:*
 187: qb_est118:*
 188: qb_est119:*
 189: qb_est120:*

190: em_qss_pln1:*
 191: em_qss_pln2:*
 192: em_qss_pro:*
 193: em_qss_rpd1:*
 194: em_qss_rpd2:*
 195: em_qss_rpd3:*
 196: em_qss_rpd4:*
 197: em_qss_rpd5:*
 198: em_qss_vrt1:*
 199: em_qss_vrt2:*
 200: em_qss_vrt3:*
 201: qb_qss1:*
 202: qb_qss2:*
 203: qb_qss3:*
 204: qb_qss4:*
 205: qb_qss5:*
 206: qb_qss6:*
 207: qb_qss7:*
 208: qb_qss8:*
 209: qb_qss9:*
 210: qb_qss10:*
 211: qb_qss11:*
 212: qb_qss12:*
 213: qb_qss13:*
 214: qb_qss14:*
 215: qb_qss15:*
 216: qb_qss16:*
 217: qb_qss17:*
 218: qb_qss18:*
 219: qb_qss19:*
 220: qb_qss20:*
 221: qb_qss21:*
 222: qb_qss22:*
 223: qb_qss23:*
 224: qb_qss24:*
 225: qb_qss25:*
 226: qb_qss26:*
 227: qb_qss27:*
 228: qb_qss28:*
 229: qb_qss29:*
 230: qb_qss30:*
 231: qb_qss31:*
 232: qb_qss32:*
 233: qb_qss33:*
 234: qb_qss34:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match Length	DB	ID	Description
1	736	39.6	757	172	BE067766
2	681.8	36.7	788	141	BE915476
3	658.6	35.4	725	21	AE528699
4	651.6	35.0	667	112	AE231122
5	650.4	35.0	706	18	AE132708
6	627.2	33.7	743	143	BE099698
7	627.2	33.7	627	121	AE911822
8	618.4	33.2	620	148	BE466445
9	618.4	33.2	666	144	BE116819
10	613	33.0	759	17	AE194859
11	604.8	32.5	704	140	BE847826
12	596	32.0	667	13	AE920102
13	591	31.8	675	17	AE169222
14	587.2	31.6	687	112	AE208542
15	583.4	31.4	807	149	BE536092
16	580	31.2	624	10	AE646678
17	569.8	30.6	957	140	BE878950
18	566.6	30.5	641	4	AE222594

486 GCGGACATGCGACATGATGCTGATCGCGGACAGTCCAGCTGTACGGAGTGGCA 545
 10b |
 241 GCGGACATGCGACATGATGCTGATCGCGGACAGTCCAGCTGTACGGAGTGGCA 300
 546 AGTGAAG 605
 10b |
 301 AGTGAAG 360
 606 AAGAGATGCGACATGATGCTGATCGCGGACAGTCCAGCTGTACGGAGTGGCA 665
 10b |
 361 AAGAGATGCGACATGATGCTGATCGCGGACAGTCCAGCTGTACGGAGTGGCA 420
 666 GCTGACATGCGACATGATGCTGATCGCGGACAGTCCAGCTGTACGGAGTGGCA 725
 10b |
 421 GCTGACATGCGACATGATGCTGATCGCGGACAGTCCAGCTGTACGGAGTGGCA 480
 726 AAGAGATGCGACATGATGCTGATCGCGGACAGTCCAGCTGTACGGAGTGGCA 785
 10b |
 481 AAGAGATGCGACATGATGCTGATCGCGGACAGTCCAGCTGTACGGAGTGGCA 540
 786 GCTGACATGCGACATGATGCTGATCGCGGACAGTCCAGCTGTACGGAGTGGCA 845
 10b |
 541 GCTGACATGCGACATGATGCTGATCGCGGACAGTCCAGCTGTACGGAGTGGCA 600
 846 TCGTGAAG 905
 10b |
 601 TCGTGAAG 660
 906 TCGTGAAG 965
 10b |
 661 TCGTGAAG 720

RESULT 5
 DEFINITION A1322708 706 bp mRNA EST 23-DEC-1998
 LOCUS m22603.y1 Soares mouse 3NDMS Mus musculus cDNA clone IMAGE:598400
 5 similar to gb:009419 Mus musculus retinoid X receptor
 interacting protein (MORF); mRNA sequence.
 ACCESSION A1322708
 VERSION A1322708.1 GI:4657147
 KEYWORDS EST
 SOURCE house mouse
 ORGANISM Mus musculus
 REFERENCES
 AUTHORS Fukuyama, M.; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;
 1 (bases 1 to 706)
 Mitter, S., Kucaba, T., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Schellander, K., Stepien, M., Tan, F., Underwood, K., Morris, R.,
 Theising, B., Wylie, L., Lennarz, G., Soares, B., Wilson, R., and
 Watson, R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Contact: Maria M/Mouse EST Project
 COMMENT Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty free through HMI, contact the
 IMAC Consortium (info@imac.hmi.gov) for further information.
 M1:364732
 This read is a RESIDUENCE of a previously sequenced mouse clone
 (correct orientation)
 putative full length read
 vector to vector length is 707
 Seq primer: -40bp from gibco
 High quality sequence step: 422.
 Location/Qualifiers
 1..706

/organism="Mus musculus"
 /strain="G57BL/6J"
 /db_xref="taxon:10090"
 /clone IMAGE:598400
 /clone_1ib="Soares mouse 3NDMS"
 /sex="male"
 /issue_type="S1000"
 /dev_stage="4 weeks"
 /lab_host="DH10B"
 /note="Vector: pT7-30-Pac (Pharmacia) with a modified
 Polylinker Site 1; Not 1 Site 2; Eco RI; 1st strand cDNA
 was primed with a Not 1 - oligo(dT) primer [5']
 3'; double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not 1 and cloned into the Not 1
 and Eco RI sites of the modified pT7 vector. RNA
 provided by Dr. Karyand Jordan, library went through
 three rounds of normalization, and was constructed by
 Bento Soares and M. Paula Bonaldo."
 BASE COUNT 144 a 230 c 184 g 148 t
 ORIGIN
 Query Match 35.0%; Score 650.4; DB 18; Length 706;
 Best Local Similarity 97.8%; Prod. No. 1.5e-142;
 Matches 681; Conservative 0; Mismatches 11; Indels 4; Gaps 2;
 1128 AATTCATCATCCGACATGCTGATCGCGGACAGTCCAGCTGTACGGAGTGGCA 1187
 1 AATTCATCATCCGACATGCTGATCGCGGACAGTCCAGCTGTACGGAGTGGCA 60
 1188 CAGAGTATGCTGATGCTGATCGCGGACAGTCCAGCTGTACGGAGTGGCA 1247
 61 CAGAGTATGCTGATGCTGATCGCGGACAGTCCAGCTGTACGGAGTGGCA 120
 1248 AATTCATCATCCGACATGCTGATCGCGGACAGTCCAGCTGTACGGAGTGGCA 1407
 121 AATTCATCATCCGACATGCTGATCGCGGACAGTCCAGCTGTACGGAGTGGCA 180
 1308 GATTCATCATCCGACATGCTGATCGCGGACAGTCCAGCTGTACGGAGTGGCA 1407
 181 GATTCATCATCCGACATGCTGATCGCGGACAGTCCAGCTGTACGGAGTGGCA 240
 1368 TCGTGAAG 1427
 241 TCGTGAAG 300
 1428 AGTGAAG 1487
 301 AGTGAAG 460
 1488 CCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1547
 361 CCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 1548 GATTCATCATCCGACATGCTGATCGCGGACAGTCCAGCTGTACGGAGTGGCA 1607
 421 GATTCATCATCCGACATGCTGATCGCGGACAGTCCAGCTGTACGGAGTGGCA 1667
 1608 GATTCATCATCCGACATGCTGATCGCGGACAGTCCAGCTGTACGGAGTGGCA 1667
 480 GATTCATCATCCGACATGCTGATCGCGGACAGTCCAGCTGTACGGAGTGGCA 1667
 1668 GATTCATCATCCGACATGCTGATCGCGGACAGTCCAGCTGTACGGAGTGGCA 1727
 540 GATTCATCATCCGACATGCTGATCGCGGACAGTCCAGCTGTACGGAGTGGCA 1787
 1728 GATTCATCATCCGACATGCTGATCGCGGACAGTCCAGCTGTACGGAGTGGCA 1847
 600 GATTCATCATCCGACATGCTGATCGCGGACAGTCCAGCTGTACGGAGTGGCA 1847
 1788 GATTCATCATCCGACATGCTGATCGCGGACAGTCCAGCTGTACGGAGTGGCA 1847

Db 657 CTCCTGCTGCTTTTATTAATAAAAAACAATAA 692

RESULT 6
LOCUS BF099698 743 bp mRNA EST 19-061-2000
DEFINITION 6017513311 NCI_TGAP_Mam1 Mus musculus cDNA clone IMAGE:3979183 5'.

ACCESSION BF099698
VERSION BF099698.1 GI:10882224
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Fukuyama; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
1 (bases 1 to 743)
NIH-MGC half 74bp pet rib seq
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
JOURNAL
AUTHORS
TITLE
COMMENT
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The J.M.A.C.L. Consortium (J.M.C.)
DNA Sequencing by: Invitrogen Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the J.M.A.C.L. Consortium/JMML at:
http://image.llnl.gov
plate: ILM9172 row: m column: 08
High quality sequence stop: 700.
Location/Qualifiers
1..743

FEATURES
Source
1..743
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/note="IMAG:3979183"
/clone_lib="NCI_TGAP_Mam1"
/issue_type="tumor, biopsy sample"
/lab_host="F1010P"
/note="Organ: mammary; Vector: pTZ19 SPOT6; Site: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: cDNA dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

BASE COUNT 156 a 235 c 213 g 138 t 1 others

ORIGIN

Query Match 33.7% Score 627.2 DB 113 Length 743.
Best Local Similarity 96.9% Proc. No. 43e-177
Matches 707; Conservation 0; Mismatches 18; Indels 9; Gaps 6

UY 683 GCTGCTGCTTTTATTAATAAAAAACAATAA 741
Db 7 GCTGCTGCTTTTATTAATAAAAAACAATAA 66
UY 742 GCTGCTGCTTTTATTAATAAAAAACAATAA 801
Db 67 GCTGCTGCTTTTATTAATAAAAAACAATAA 126
UY 802 GCTGCTGCTTTTATTAATAAAAAACAATAA 861
Db 127 GCTGCTGCTTTTATTAATAAAAAACAATAA 186
UY 862 GCTGCTGCTTTTATTAATAAAAAACAATAA 921
Db 187 GCTGCTGCTTTTATTAATAAAAAACAATAA 246
UY 922 GCTGCTGCTTTTATTAATAAAAAACAATAA 981
Db 247 GCTGCTGCTTTTATTAATAAAAAACAATAA 306

UY 982 GTACATGCGGCTGCTTTTATTAATAAAAAACAATAA 1041
Db 407 GTACATGCGGCTGCTTTTATTAATAAAAAACAATAA 366
UY 1042 GTACATGCGGCTGCTTTTATTAATAAAAAACAATAA 1101
Db 367 GTACATGCGGCTGCTTTTATTAATAAAAAACAATAA 426
UY 1102 GTACATGCGGCTGCTTTTATTAATAAAAAACAATAA 1161
Db 427 GTACATGCGGCTGCTTTTATTAATAAAAAACAATAA 486
UY 1162 GTACATGCGGCTGCTTTTATTAATAAAAAACAATAA 1221
Db 487 GTACATGCGGCTGCTTTTATTAATAAAAAACAATAA 546
UY 1222 GTACATGCGGCTGCTTTTATTAATAAAAAACAATAA 1281
Db 547 GTACATGCGGCTGCTTTTATTAATAAAAAACAATAA 606
UY 1282 GTACATGCGGCTGCTTTTATTAATAAAAAACAATAA 1342
Db 606 GTACATGCGGCTGCTTTTATTAATAAAAAACAATAA 662
UY 1342 GTACATGCGGCTGCTTTTATTAATAAAAAACAATAA 1401
Db 663 GTACATGCGGCTGCTTTTATTAATAAAAAACAATAA 718
UY 1402 GTACATGCGGCTGCTTTTATTAATAAAAAACAATAA 1461
Db 719 GTACATGCGGCTGCTTTTATTAATAAAAAACAATAA 772

RESULT 7
AW911822
DEFINITION
633612191 STARS_Eukaryote-31341-21000 Mus musculus cDNA clone
IMAGE:151750 5' similar to GI:509419 Mus musculus retina X
receptor interacting protein (MOUSE);, mRNA sequence.
AW911822
AW911822.1 GI:8577405
EST.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
house mouse.
Mus musculus
Fukuyama; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 627)
NCI-TGAP http://www.ncbi.nlm.nih.gov/tcgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496 1550
Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through J.M.C.; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40bp from cDNA
High quality sequence stop: 477.
Location/Qualifiers
1..627

FEATURES
Source
1..627
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_lib="NCI_TGAP_Mam1"
/issue_type="tumor, biopsy sample"
/lab_host="F1010P"
/note="Vector: pTZ19-SP6 (Pharmacia) with a modified
polylinker. 3' strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
ligated to Eco RI adaptors (Pharmacia), digested with Not


```

/strain "C57H1."
/db_xref "taxon:10090"
/clone "IMAGE:1886535"
/clone_lib "Suidae mouse liver mlia"
/sex "female"
/dev_stage "adult"
/lab_host "DH108"
/notes "Jordan; Vector: pME18S-FL3; Site_1: DraIII
(CACCTGCTG); Site_2: DraIII (CACCATGCTG); 1st strand cDNA
was primed with an oligo(dT) primer
(AATGACCTCTTTTCTTTTCTTTT); double stranded cDNA was
ligated to a DraIII adaptor (TCTCCGCTACCTG), digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACCTGCTG, 3' site CACCATGCTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Samu Suidae (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTCTGCTCTCAAAAGCTGCG and 3' end
primer CGACCTGACCTGAGACAA."
BASE COUNT      155 a      241 g      167 t      5 others
ORIGIN
Query Match      43.08; Score 613; DB 17; Length 759;
Host Local Similarity 92.48; Pred. No. 9,40-134;
Matches 709; Conservative 0; Mismatches 50; Indels 4; Gaps 6;

QY 1083 TCACCTAATACAAACATTCGACCTTACCTTTCAGCTGCAATTCATCAATCA 1142
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 759 TCACCTAATACAAACGCG-CTTCACCTGCTCAAGCT--CGACGTAAATTCATCAATCA 703

QY 1143 TCTTCATCTTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1202
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 702 TCTTCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 647

QY 1203 TTAATTCATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1262
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 646 TTAATTCATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 589

QY 1263 AGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1322
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 587 AGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 528

QY 1323 AGCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1382
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 527 AGCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 468

QY 1383 CCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1442
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 467 CCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 408

QY 1443 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1502
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 407 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 348

QY 1503 GCTCTCTGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 1562
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 347 GCTCTCTGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 289

QY 1563 CAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1622
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 287 CAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 229

QY 1623 CCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1682
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 228 CCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 169

QY 1683 TGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1742
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 168 TGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 109

QY 1743 GTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1802

```

```

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 108 CTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 49

QY 1803 TTTAATAAAAAAATAAATAAATTGACTGATCAATGCTCAAAAAA 1850
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 48 TTTAATAAAAAAATAAATAAATTGACTGATCAATGCTCAAAAAA 1

RESULT 11
BE847826 704 bp mRNA 26-SEP-2000
LOCUS 2641-07 y1 sources_2BMT Mus musculus cDNA clone
DEFINITION IMAGE:3419268 5' similar to gb:U09419 Mus musculus retinoid X
receptor interacting protein (MURF).; mRNA sequence.
ACCESSION BE847826
VERSION BE847826.1 GI:10306165
KEYWORDS TEST.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 704)
AUTHORS Mammalia; Eutheria; Rodentia; Scurionathia; Muridae; Mus.
TITLE NCI-CCAP http://www.ncbi.nlm.nih.gov/nciccap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CCAP),
COMMENT Unpublished (1997)
CONTACT: Robert Strausberg, Ph.D.
TEL: (301) 496-1550
Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through UNL; contact the
IMAGE Consortium (info@image.intl.gov) for further information.
MGI:1095080
Seq primer: -40RP from Gibco
High quality sequence stop: 463.
Location/Qualifiers
1..704
/organism="Mus musculus"
/strain="C57H1/6J"
/db_xref="taxon:10090"
/clone="IMAGE:3419268"
/clone_lib "sources_thymus_2BMT"
/sex="male"
/tissue_type="thymus"
/lab_host="DH108"
/dev_stage="4 weeks"
/notes="Vector: pT73D-pac (Pharmacia) with a modified
poly(A) tail; Not 1: Site_2: Eco RI; 1st strand cDNA
was primed with a Not 1 oligo(dT) primer (5'
TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
3'); double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not 1 and cloned into the Not 1
provided by Dr. Bertrand Jordan. Library went through two
rounds of normalization, and was constructed by Bento
Scores and M. Fatima Ronaldi."
BASE COUNT      144 a      237 c      182 g      140 t      1 others
ORIGIN
Query Match      32.58; Score 604.8; DB 140; Length 704;
Host Local Similarity 96.58; Pred. No. 7,96-122;
Matches 682; Conservative 0; Mismatches 18; Indels 7; Gaps 6;

QY 1029 AATAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1086
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2 AATAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 61

QY 1089 AATAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1148
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 62 AATAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 121

QY 1149 AGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1208
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```


|||||
690 AAACGATTCAGAAATGAGAAACGAGAGCAACACCCGATTTGAGG 738

Search completed: April 29, 2001, 15:34 25
Job time: 4738 sec

[illegible]


```

01-MAY-2000 (TEMBREL, 13, last sequence update)
01-OCT-2000 (TEMBREL, 15, last annotation update)
DE BRUSHONE RECEPTOR ISOFORM A.
GN ECR.
OS Choristoneura fumiferana (Spruce budworm).
OC Eukaryota; Metazoa; Artropoda; Tracheata; Hexapoda; Insecta;
OC Protostoma; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditystia;
OC Tortricidae; Tortricinae; Tortricidae; Choristoneura.
OX NBI_Taxid 71411.
RN [11]
RP SEQUENCE FROM N.A.
RA Kotaballi R., Ladd T.R., Bhadialla I.S., Krell P.J., Solt S.S.,
RA Retnakaran A., Palli S.R.;
RT "Studies of two polypeptide receptor isoforms of the spruce budworm,
RT Choristoneura fumiferana."
RT Submitted (2001-1999) to the EMBL/Genbank/Tras databases.
OC EMBL: AF092030; AAC61596.27.
DE HSSP: P20394; IABY.
DE INTERPRO: IPR000324.
DE INTERPRO: IPR000546.
DE INTERPRO: IPR001723.
DE INTERPRO: IPR004069.
DE PFAM: PF00104; Hormone_recep_1.
DE PFAM: PF00105; Z1-C4; 1.
DE PRINTS: PR00047; STROHORNKR.
DE PRINTS: PR00350; VITAMINOR.
DE PRINTS: PR01283; ECDYSTEROIDR.
DE PRINTS: PS00031; NOCTURNAL_RECEPTOR_1.
KW Receptor; Transcription regulation; DNA binding; Nuclear protein;
KW Zinc finger.
SC SEQUENCE: 514 AA; 57621 MW; F110540A30305602 cfr64.

Query Match 41.8%; Score 736; DB 5; Length 513;
Host Local Similarity 46.4%; Pred. No. 1 50-54;
Matches 172; Conservative 84; Mismatches 168; Indels 48; Gaps 11;

01-MAY-2000 (TEMBREL, 13, last sequence update)
01-OCT-2000 (TEMBREL, 15, last annotation update)
DE BRUSHONE RECEPTOR ISOFORM A.
GN ECR.
OS Choristoneura fumiferana (Spruce budworm).
OC Eukaryota; Metazoa; Artropoda; Tracheata; Hexapoda; Insecta;
OC Protostoma; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditystia;
OC Tortricidae; Tortricinae; Tortricidae; Choristoneura.
OX NBI_Taxid 71411.
RN [11]
RP SEQUENCE FROM N.A.
RA Kotaballi R., Ladd T.R., Bhadialla I.S., Krell P.J., Solt S.S.,
RA Retnakaran A., Palli S.R.;
RT "Studies of two polypeptide receptor isoforms of the spruce budworm,
RT Choristoneura fumiferana."
RT Submitted (2001-1999) to the EMBL/Genbank/Tras databases.
OC EMBL: AF092030; AAC61596.27.
DE HSSP: P20394; IABY.
DE INTERPRO: IPR000324.
DE INTERPRO: IPR000546.
DE INTERPRO: IPR001723.
DE INTERPRO: IPR004069.
DE PFAM: PF00104; Hormone_recep_1.
DE PFAM: PF00105; Z1-C4; 1.
DE PRINTS: PR00047; STROHORNKR.
DE PRINTS: PR00350; VITAMINOR.
DE PRINTS: PR01283; ECDYSTEROIDR.
DE PRINTS: PS00031; NOCTURNAL_RECEPTOR_1.
KW Receptor; Transcription regulation; DNA binding; Nuclear protein;
KW Zinc finger.
SC SEQUENCE: 514 AA; 57621 MW; F110540A30305602 cfr64.

Query Match 41.8%; Score 736; DB 5; Length 513;
Host Local Similarity 46.4%; Pred. No. 1 50-54;
Matches 172; Conservative 84; Mismatches 168; Indels 48; Gaps 11;

```

```

RESULT 10
077255 ID 077255 PRELIMINARY: PRT: 541 AA.
AC 077255:
DT 01-NOV-1998 (TEMBREL, 08, created)
DT 01-MAY-2000 (TEMBREL, 13, last sequence update)
DT 01-OCT-2000 (TEMBREL, 15, last annotation update)
DE ECDYSTEROID RECEPTOR ECR-B.
OS Choristoneura fumiferana (Spruce budworm).
OC Eukaryota; Metazoa; Artropoda; Tracheata; Hexapoda; Insecta;
OC Protostoma; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditystia;
OC Tortricidae; Tortricinae; Tortricidae; Choristoneura.
OX NBI_Taxid 71411.
RN [11]
RP SEQUENCE FROM N.A.
RA Kotaballi R., Ladd T.R., Bhadialla I.S., Krell P.J., Solt S.S.,
RA Retnakaran A., Palli S.R.;
RT "Cloning and developmental expression of the polypeptide receptor gene
RT from the spruce budworm, Choristoneura fumiferana."
RT Dev. Genet. 17:319-330(1995).
[2]
RN [12]
RP SEQUENCE FROM N.A.
RA Kotaballi R., Palli S.R., Ladd T.R., Retnakaran A.;
RA Submitted (2001-1998) to the EMBL/Genbank/Tras databases.
SC EMBL: 029531; AAC36491.27.
DE HSSP: P20394; IABY.
DE INTERPRO: IPR000324.
DE INTERPRO: IPR000536.
DE INTERPRO: IPR001629.
DE INTERPRO: IPR001723.
DE INTERPRO: IPR003689.
DE PFAM: PF00104; Hormone_recep_1.
DE PFAM: PF00105; Z1-C4; 1.
DE PRINTS: PR00047; STROHORNKR.
DE PRINTS: PR00350; VITAMINOR.
DE PRINTS: PR00398; STROHORNKR.
DE PRINTS: PR01283; ECDYSTEROIDR.
DE PRINTS: PS00031; NOCTURNAL_RECEPTOR_1.
KW Receptor; Transcription regulation; DNA binding; Nuclear protein;
KW Zinc finger.
SC SEQUENCE: 541 AA; 61152 MW; C78D28F7D2B688CD cfr64.

Query Match 41.5%; Score 728.5; DB 5; Length 541;
Host Local Similarity 46.9%; Pred. No. 6 80-84;
Matches 168; Conservative 81; Mismatches 166; Indels 45; Gaps 10;

01-MAY-2000 (TEMBREL, 13, last sequence update)
01-OCT-2000 (TEMBREL, 15, last annotation update)
DE BRUSHONE RECEPTOR ISOFORM A.
GN ECR.
OS Choristoneura fumiferana (Spruce budworm).
OC Eukaryota; Metazoa; Artropoda; Tracheata; Hexapoda; Insecta;
OC Protostoma; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditystia;
OC Tortricidae; Tortricinae; Tortricidae; Choristoneura.
OX NBI_Taxid 71411.
RN [11]
RP SEQUENCE FROM N.A.
RA Kotaballi R., Ladd T.R., Bhadialla I.S., Krell P.J., Solt S.S.,
RA Retnakaran A., Palli S.R.;
RT "Studies of two polypeptide receptor isoforms of the spruce budworm,
RT Choristoneura fumiferana."
RT Submitted (2001-1999) to the EMBL/Genbank/Tras databases.
OC EMBL: AF092030; AAC61596.27.
DE HSSP: P20394; IABY.
DE INTERPRO: IPR000324.
DE INTERPRO: IPR000546.
DE INTERPRO: IPR001723.
DE INTERPRO: IPR004069.
DE PFAM: PF00104; Hormone_recep_1.
DE PFAM: PF00105; Z1-C4; 1.
DE PRINTS: PR00047; STROHORNKR.
DE PRINTS: PR00350; VITAMINOR.
DE PRINTS: PR01283; ECDYSTEROIDR.
DE PRINTS: PS00031; NOCTURNAL_RECEPTOR_1.
KW Receptor; Transcription regulation; DNA binding; Nuclear protein;
KW Zinc finger.
SC SEQUENCE: 514 AA; 57621 MW; F110540A30305602 cfr64.

Query Match 41.5%; Score 728.5; DB 5; Length 541;
Host Local Similarity 46.9%; Pred. No. 6 80-84;
Matches 168; Conservative 81; Mismatches 166; Indels 45; Gaps 10;

```

[illegible]

```

PRINTS: PR00047: SEQUENCER.
PROTEIN: P00047: NUCLEAR RECEPTOR 1.
RECEPTOR: Transcription regulation; DNA-binding; Nuclear protein;
Zinc-finger.
FT DN_BIND 78 145 NUCLEAR RECEPTOR-TYPE.
FT ZN_FING 78 96 C4-TYPE.
FT ZN_FING 116 140 C4-TYPE.
FT DOMAIN 216 446 LIGAND BINDING (POTENTIAL).
FT DOMAIN 162 168 POLY-GIN.
FT DOMAIN 169 172 POLY-PRO.
SEQUENCE 446 AA: 49719 MW: 7415463540200DF CR064)

Query Match 98.88: Score 23097; DB 1: Length 446;
Best Local Similarity 99.88: Prod. No. 9, Zc-156;
Matches 445: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

UY 1 MSSPTSSLDTPVNGSGSPSTASPTKPGVPTTPPSSSSAVTVLLEPPDP 60
DB 1 MSSPTSSLDTPVNGSGSPSTASPTKPGVPTTPPSSSSAVTVLLEPPDP 60
UY 61 EPRKPKTAPPMKCHELQVWQKASGFHYNVLSGVYKGFPRSSVVRZACVYA 120
DB 61 EPRKPKTAPPMKCHELQVWQKASGFHYNVLSGVYKGFPRSSVVRZACVYA 120
UY 121 TQGMATMPKPCGLPRKPCFVACMPGVVSEFGKPPFQVQVQPPSPFAVSS 180
DB 121 TQGMATMPKPCGLPRKPCFVACMPGVVSEFGKPPFQVQVQPPSPFAVSS 180
UY 181 GHPAASPTSSASSSGSGEGGGLTAACELMUGVAAALQGNKSFSGVPTWPLG 240
DB 181 GHPAASPTSSASSSGSGEGGGLTAACELMUGVAAALQGNKSFSGVPTWPLG 240
UY 241 ADPQSDAKQOORFAHFELALISVQFIVDPKQVPSFGGPPFALIKASTTEMLL 300
DB 241 ADPQSDAKQOORFAHFELALISVQFIVDPKQVPSFGGPPFALIKASTTEMLL 300
UY 301 TAPVNIETETITKKTATYSKQPHKALQVETINLFFSSAMRKLGLDAVALIA 360
DB 301 TAPVNIETETITKKTATYSKQPHKALQVETINLFFSSAMRKLGLDAVALIA 360
UY 401 INFPSADPQNGQSRVVALQGVVVALISTYRKRQVGLPFRPMKLVSLTSSVH 420
DB 401 INFPSADPQNGQSRVVALQGVVVALISTYRKRQVGLPFRPMKLVSLTSSVH 420
UY 421 SFQVPAKLQKRLPGLLSLFWVHE 446
DB 421 SFQVPAKLQKRLPGLLSLFWVHE 446

RESULT 2
NR02_RAT STANALDO. PRT. 446 AA.
AC 064755: Q62694:
DB 40 MAY-2000 (Ref. 49, Cited)
DB 40 MAY-2000 (Ref. 49, Last sequence update)
DB 01 OCT-2000 (Ref. 40, Last annotation update)
DB OXSTEROLS RECEPTOR LXR BETA (LIVER X RECEPTOR BETA) (NUCLEAR ORPHAN
RECEPTOR LXR BETA) (CHOLESTEROL-EXRESSED NUCLEAR RECEPTOR) (OR)
DB (ORPHAN NUCLEAR RECEPTOR OR-1).
CN NP02 OF LXR.
CS KATUUS NORVADICUS (RAT).
CS EUCALYPTUS MEGALOCALYPTUS (Eucalyptus).
CS MAMMALIA: EUTHALIA: RODENTIA: SCIRODONTIA: MURIDAE: MURINAE: RATIUS.
CS NUCLEOTIDE: POLY-
CN 11.
RP SEQUENCE FROM N.A.
RC STRAIN: SPRAGUE-DAWLEY: TISSUE: Liver.
KA MEDLINE: 95179298: PubMed 7892303.
KA Takeda M., Hamada F., Ito Y., Wikstrom A.O., Pelto-Pulkko M.,
KA Gustafsson J.-A.:
KA 7081: a member of the nuclear receptor superfamily that interacts
KA with the 9-cis-retinoic acid receptor."

```

```

PL Proc. Natl. Acad. Sci. U.S.A. 92:2096-2100(1995).
RN 121.
RP SEQUENCE FROM N.A.
RC STRAIN: SPRAGUE-DAWLEY: TISSUE: Vagina.
KA MEDLINE: 95062154: PubMed 7971966.
FA Saito C., Kokubo J.M., Eitjaka R.A., Iino S.:
RT "Orphan" receptor: a receptor that modulates gene activation by
RT retinoic acid and thyroid hormone receptors."
PL Proc. Natl. Acad. Sci. U.S.A. 91:10809-10814(1994).
CC 1.
CC ORPHAN RECEPTOR. KINOS DIFFERENTIALLY TO OTHER STRAINED
CC OXSTEROLS DIRECT REPEATS HAVING THE CONSENSUS HALF-SITE.
CC SEQUENCE 57-AGGTCA-3' AND 4-NT SPACING (OR 4).
CC -1 SUBUNIT: FORMS A DETERMINER WITH RXR.
CC -1 SUBUNIT: LIGATION: NUCLEAR (POTENTIAL).
CC -1 SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
CC NR SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL database.
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usurped by, and for commercial
CC entities requires a license agreement with the EMBL database.
CC or send an email to license@ebi.ac.uk.
CC
CC EMBL: 029789; AAA59522.1;
CC EMBL: 014533; AAA52461.1;
CC HSRP: P03372; IBCP.
CC
CC InterPro: IPR000536;
CC InterPro: IPR001628;
CC Pfam: PF00104; hormone_rec_1.
CC Pfam: PF00105; z1-c4_1.
DB PRINTS: PR00047: SEQUENCER.
DB PROSITE: P00047: NUCLEAR RECEPTOR 1.
KM Receptor: Transcription regulation; DNA-binding; Nuclear protein;
FM Zinc-finger.
FT DN_BIND 78 145 NUCLEAR RECEPTOR-TYPE.
FT ZN_FING 78 98 C4-TYPE.
FT ZN_FING 116 140 C4-TYPE.
FT DOMAIN 216 446 LIGAND BINDING (POTENTIAL).
FT DOMAIN 162 168 POLY-GIN.
FT DOMAIN 169 172 POLY-PRO.
FT COMPLET 33 33 G--V (IN REF. 2).
FT COMPLET 52 54 MISSING (IN REF. 2).
FT COMPLET 219 219 A--V (IN REF. 2).
SEQUENCE 446 AA: 49735 MW: 1406042705FAD CR064)

Query Match 98.88: Score 2284; DB 1: Length 446;
Best Local Similarity 98.28: Prod. No. 5, Zc-154;
Matches 448: Conservative 6: Mismatches 2: Indels 0: Gaps 0:

UY 1 MSSPTSSLDTPVNGSGSPSTASPTKPGVPTTPPSSSSAVTVLLEPPDP 60
DB 1 MSSPTSSLDTPVNGSGSPSTASPTKPGVPTTPPSSSSAVTVLLEPPDP 60
UY 61 EPRKPKTAPPMKCHELQVWQKASGFHYNVLSGVYKGFPRSSVVRZACVYA 120
DB 61 EPRKPKTAPPMKCHELQVWQKASGFHYNVLSGVYKGFPRSSVVRZACVYA 120
UY 121 TQGMATMPKPCGLPRKPCFVACMPGVVSEFGKPPFQVQVQPPSPFAVSS 180
DB 121 TQGMATMPKPCGLPRKPCFVACMPGVVSEFGKPPFQVQVQPPSPFAVSS 180
UY 181 GHPAASPTSSASSSGSGEGGGLTAACELMUGVAAALQGNKSFSGVPTWPLG 240
DB 181 GHPAASPTSSASSSGSGEGGGLTAACELMUGVAAALQGNKSFSGVPTWPLG 240
UY 241 ADPQSDAKQOORFAHFELALISVQFIVDPKQVPSFGGPPFALIKASTTEMLL 300
DB 241 ADPQSDAKQOORFAHFELALISVQFIVDPKQVPSFGGPPFALIKASTTEMLL 300
UY 301 TAPVNIETETITKKTATYSKQPHKALQVETINLFFSSAMRKLGLDAVALIA 360

```

[illegible][illegible]

1

2

3

08-09 365-576-3 x MM009419

Align seq 1/1 for MM009419 from: 1 to: 1841

```

1 MetSerSerProThrSerSerLeuAspThrProValProCysAsnGlySer 17
  |||
141 ATGCTTGGCCCAATCAACTTCTCTGAAATCACTGGCTGGCAAGTAAGTCTG 180
  |||
17 ProGluProSerThrSerAlaThrSerProThrLeuGlySerGluGlyG 34
  |||
181 TCTTCAAGCTCACTACCTGCTGACCACTATTAAAGCAAGACAGAGAG 230
  |||
34 LeuGluHisAspProProLeuGlySerGluGlySerSerSerAlaIle 50
  |||
241 AGCAATCTGACACTCTCTCAAGCTTTGAAGGCTTCAGCTTCTGCTTAAATC 280
  |||
51 ValValIleLeuGluProGluAspGluProGluValArgGlyGlySer 67
  |||
281 GTGCTTATCTTAAAGCTCAATCAATCACTGACCTGACCTCAAGACAGAAAG 330
  |||
67 yProGluLeuProGlySerMetGluGlyHisGluGlySerArgValGlySer 84
  |||
341 GCGAGATCCGAGAGATGCTGAGCCATGACCTGTCGCGCTGCTGGAGACA 380
  |||
84 yValAspSerGlyProHisSerGlyAsnValLeuSerGlySerGluGlyGly 100
  |||
381 AGCTTTCGAGCTTCCATCACTCAACGCTGTCAGCTGCGAAGGCTCCCAAGGC 430
  |||
101 ThrGluAlaGluGluSerValValHisSerGlyGlyValArgGlyVal 117
  |||
431 TCTTTCGAGAGAGATGCTGAGCTCAATCAATCACTGAGCTGCTGCTG 480
  |||
117 GeliSerGlyThrCysGluMetAspAlaPheMetArgArgGlySerGlu 134
  |||
481 GCGATAGGAGAACTTGCATGATGGAATGCTTCACTGGCGCTCAAGTGTAG 530
  |||
144 GueGlyAlaAlaArgGlySerGlySerGluAlaGluGlyMetArgGlu 150
  |||
541 TCTGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 580
  |||
151 LeuSerGluGluGluGluGluGluGlySerArgGluGluGlySerGlu 167
  |||
581 GCTTTCGAGAGAGATGCTGAGCTCAATCAATCACTGAGCTGCTGCTG 630
  |||
167 GeliProProProProSerGluProAlaAlaSerSerSerGlyArgPro 184
  |||
641 GATGCTCAACCAATCACTCACTGAGCTGAGCTGAGCTGAGCTGAGCTG 680
  |||
184 LeuAlaSerProSerGlySerSerGluAlaSerSerSerGlySerGly 200
  |||
681 GCGATAGGAGAACTTGCATGATGGAATGCTTCACTGGCGCTCAAGTGTAG 730
  |||
201 GluGlyIleGluGluSerGluAlaAlaGluGluGluGluGluGluGlu 217
  |||
741 GAGGATATGCACTTACCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 780
  |||
217 AlaAlaAlaGluGluGluGlySerGlySerGlySerGlySerGlySer 234
  |||
781 TCGCTGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 830
  |||
244 GThrProThrProLeuGluGlyAlaAspProGluSerArgAspAlaArg 250
  |||
841 TCAATGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 890
  |||
261 GluArgPheAlaHisPheThrGluGluAlaIleLeuSerValGluGlu 267
  |||
881 CAAGCTTTTGGCTGCTTCAATCACTGAGCTTACCTGATCTGCTGCTGAG 930
  |||
267 GValAspPheAlaGlySerGlyValProGlyThrLeuGluGluGlyVal 284
  |||
941 TGTGCACTTTCATCAAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 980
  |||
284 GProGluAlaAlaGluGlySerThrLeuGluGluGluGluGluGlu 300

```

```

981 ACCAGATGCGCTCTCTGAGAGGCTGCACTGATGATGATGATGATGATG 1030
  |||
1031 ThrAlaArgArgTyrAsnHisGluThrTrpGlyIleThrLeuGlySer 317
  |||
1041 ACAGCAGACAGCTTACACACGACAGACGACGACGACGACGACGACGAC 1080
  |||
317 PheThrTyrSerGlySerAspAspPheHisArgAlaGluGlyGluValGlu 334
  |||
1081 CTTCACTGACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1130
  |||
334 IleLeuAspProIlePheGluIlePheSerArgAlaMetArgAlaGluGly 350
  |||
1131 TCATCAATGCACTCTGCAATCTTCTGCGGCGCATGCGGCGCTGCGCTG 1180
  |||
351 AspAspAlaGluGlyArgAlaGluGluGluGluGluGluGluGluGlu 367
  |||
1181 GACCATGCAAGCATGAGCTCTGCTTATGCGCATCAACATCTCTGCAAC 1230
  |||
367 PArgProSerValGluGluProSerArgValGluAlaGluGluGluGlu 384
  |||
1231 TCGCTTATGCTGACAGACCTGACCTGCTGCAAGCTCTGCTGCTGCTG 1280
  |||
384 yValGluAlaGluGluSerGlyThrArgGlySerArgProGluAspGlu 400
  |||
1281 ACCTGAGAGCGCTCTCTGCTACAGAGATCAAGCGCGCACAGACAGCA 1330
  |||
401 LeuArgPheProArgMetLeuMetGlySerValSerLeuArgGluThr 417
  |||
1331 CTCTGCTTCTGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
  |||
417 PserValHisSerGluGluValPheAlaAlaArgAlaGluGluAspGly 434
  |||
1381 CTGCGTGCAGCTGCGAGCGAGCTTTCATGCGCAGCTGCAATCAATCA 1430
  |||
434 GueProProLeuSerGluIleTrpAspValHisGlu 446
  |||
1431 TCGCGTCTCTGCTGCGCAATCTGCAATGCTGCAATGCTGCAATG 1466
  |||
seq_name: gb_f02:RNU20489

```

```

seq_documentation_block:
  Locus: RNU20389, 1877 bp, mRNA
  DEFINITION: Ratius norvegicus orphan nuclear receptor OR-1 mRNA, complete cds.
  ACCESSION: U20389
  VERSION: 020389.1 GI:665941
  KEYWORDS:
  SOURCE:
  ORGANISM:
    Rat.
    Norvegicus
    Eukaryota; Metazoa; Chordata; Gnathostomata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
    Rattus.
  REFERENCE:
    1 (bases 1 to 1877)
    Teboul M., Emmerk E., Li G., Wikstrom A.C., Petro-Bukko M. and
    Gustafsson J.A.
    OR-1, a member of the nuclear receptor superfamily that interacts
    with the 9-cis-retinoic acid receptor
    Proc. Natl. Acad. Sci. U.S.A. 92 (6), 2096-2100 (1995)
    56199298
  REFERENCE:
    2 (bases 1 to 1877)
    Emmerk E.
    Direct Submission
    Submitted (40-JAN-1995) EMBL
    Novum, Center for Biotechnology, Karolinska Institute, Novum,
    Huddinge, S-141 86, Sweden
    Location/Qualifiers
    1..1877
    /organism="Rattus norvegicus"
    /strain="Sprague-Dawley"
    /db_xref="taxon:10116"
    /clone_lib="Jambaa ZAP II cDNA library, Stratagene catalog
    number 946, 514"
    /sex="male"

```

```

FEATURES
  source

```


SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 1934)
AUTHORS Emmert, E. and Gustafsson, J. A.
TITLE OR-1 ON OPIAN RECEPTOR BELONGING TO THE NUCLEAR RECEPTOR FAMILY
JOURNAL Patent: WO 9605360-A 1 22-FEB-1996;
KAROHIO AH (SE)
COMMENT other publication AU 4484495 960307
FEATURES
Location/Qualifiers
Source 1..1934
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/tissue_type="LIVER"
NAME COUNT 440 a 988 c 544 g 462 t
ORIGIN
alignment_scores:
quality: 2284.00 length: 446
ratio: 5.121 gaps: 0
Percent Similarity: 100.000 Percent Identity: 98.206
alignment_block:
US 09-365-576-3 x AA9189 ..
Align seq 1/1 to: AA9189 from: 1 to: 1934
1 Met Ser Ser Pro Thr Ser Ser Leu Asp Thr Pro Val Pro Gly Val Ser Gly 17
191 AATGCTTCGCGCAACAGTCTCTGACACATCTCTGGCTGGGAATGGCTTC 240
17 Thr Cys Thr Pro Thr Ser Ala Thr Ser Pro Thr Thr Cys Ser Gly Val 34
241 TCTTCAACATCACTACTTCTCTCACTTCAACCACTATTAAGAGGAGGAGC 290
34 Thr Thr Thr Ser Pro Thr Pro Gly Ser Gly Ser Ser Ala Thr Ile 50
291 AAGAAAGTCAATCACTTCAAGCTCTGAAGGCTTCACTTCCCTACATC 340
51 Val Val Ile Leu Glu Thr Glu Asp Thr Pro Glu Val Cys Val Cys 67
341 GTGGTATCTTAAAGCAAGACATCAAACTGAGCGCAAGCGCAAGAAAGG 390
57 Thr Pro Ala Thr Ser Met Thr Cys Thr His Cys Leu Cys Ser Val Cys 84
491 TCTGAGAGGAAATATCTGAGATATGAGATGATGAGAGGAGGAGGAGCA 440
84 Val Ala Ser Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr 100
441 AAGCTTCGCGCTTCCTCAACAAATGCTTCACTTCAAGGCTTCAAGAAAG 490
101 Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr 117
491 TCTTTCGAGTAACTGCTCACTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 540
117 Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr 134
541 GAGTAAAGGAGTCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 590
134 Cys Val Ala Leu Val Cys Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr 150
591 TCTTTCGAGTAACTGCTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 640
151 Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr 167
641 CTTTTCGAGGAGTAACTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 690
167 Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr 184
691 GAGTAAAGGAGTCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 740

184 LALASCTTCGCTTCCTGAGGAGTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 200
741 CAGCTTCGCTTCGCTTCAAGGAGTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 790
201 Glu Gly Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr 217
791 GAGGCTTCAAGGAGTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 840
217 LALASCTTCGCTTCCTGAGGAGTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 244
841 TCGCGCTTCAAGGAGTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 890
234 Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr 250
891 TCGCGCTTCAAGGAGTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 940
251 Glu Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr 267
941 CAGGCTTCAAGGAGTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 990
267 Cys Val Ala Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr 284
991 CCGTCACTTCGCGCAAGGAGTCAAGGCTTCAAGGCTTCAAGGCTTCAAGG 1040
284 Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr 300
1041 ATGAGATGCTTCCTGAGGAGTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 1090
301 Thr Ala Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr 317
1091 ATGAGGAGGAGTCAAGGAGTCAAGGCTTCAAGGCTTCAAGGCTTCAAGG 1140
317 Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr 344
1141 CTTCACTTCAAGGAGTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 1190
344 Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr 350
1191 TCAATGAGGAGTCAAGGAGTCAAGGCTTCAAGGCTTCAAGGCTTCAAGG 1240
351 Asp Asp Ala Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr 367
1241 GAGGAGTCAAGGAGTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 1290
367 Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr 384
1291 CCGGCTTCAAGGAGTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 1440
384 Val Val Glu Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr 400
1341 ATGAGGAGGAGTCAAGGAGTCAAGGCTTCAAGGCTTCAAGGCTTCAAGG 1490
401 Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr 417
1491 CCGGCTTCAAGGAGTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 1540
417 Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr 444
1441 CCGGCTTCAAGGAGTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 1490
444 Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr 466
1491 TCGGCTTCAAGGAGTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 1540

seq_name: gb_pat2:146766

seq_documentation_block:

LOCUS 146766 1959 bp DNA

DEFINITION Sequence 3 from patent US 5,696,15.

ACCESSION 146766

VERSION 146766.1 GI:2470741

KEYWORDS

FBI 07-OCT-1997

SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1959)
 AUTHORS Liou, S. and Song, C.
 TITLE Isolated nucleic acid encoding a ubiquitous nuclear receptor
 JOURNAL Patent US 5636616-A x 17-JUN-1997
 FEATURES
 source location/qualifiers
 1..1959
 BASE COUNT 406 a 608 c 564 g 481 t
 ORIGIN
 alignment_scores:
 quality: 2247.50 length: 446
 Ratio: 5.085 gaps: 1
 Percent Similarity: 99.103 Percent Identity: 97.085
 alignment_block:
 08-09-365-576-3 x 146766 ..
 Align seq 1/1 to: 146766 from: 1 to: 1959
 1 MetSerSerProProHisSerSorteAspHisProValProGlyAsnGlySer 17
 256 ATGCTTCCTCCCAACAGTCTCTGGACACTCTCTCTGCAATGCTGCAATGCTG 405
 17 ProGlyProSerTheSerAlaHisSerProHisTheGlySerGlyGlyGly 44
 406 TGGCAGTCCAGCACTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 455
 34 InsulinAspProProProGlySerGlyGlyGlySerSorteAlaTyrPhe 50
 356 ASGAGATGATGACCTCCAGAGCTCTGCAAGGAGGAGGAGGAGGAGGAGGAGG 405
 51 ValValHisLeuGlyHisProGlyHisSerHisProGlyHisValGlySerGly 67
 406 GTG.....CAATGCAAGGATGAACTGAGAGGCAATGAGCAAGGCAAGGCA 446
 67 TyrAlaProGlyMetLeuGlyHisGlyGlyLeuGlySerValGlyGlyAsp 84
 447 TGGGGGCGGCAAGAGCTGAGGAGCACTGAGCTGAGCTGAGCTGAGGAGCA 496
 84 ValAlaSerGlyPheHisTyrAspValLeuSerGlySerGlyGlyGlyGly 100
 497 AGGCTCTGGGCTTCCACTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 546
 101 PhePheAspArgSerValValHisGlyGlyAlaGlyAlaGlyTyrAlaGlySer 117
 547 TTCTTCTGGGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 596
 117 GglySerGlyThyGlySerMetAspAlaPheMetArgAlaGlySerGlyGly 134
 597 GGGCAAGGCAAGCTGGCAGATGATGCTTCAAGGCGGCAATGCTGCTGCTG 646
 134 GglySerGlyLeuArgTyrSerGlySerGlyAlaGlyGlyGlyGlyGlyVal 150
 647 TCTGTAAAGTGGCAAGTGGCAAGGAGCTGGTAAAGGCAAGGCAAGTGGT 696
 151 LeuSerGlyMetGlyHisLeuAlaGlySerGlyGlyGlyGlyGlyGlyGly 167
 697 CTTCCTCAACCAACATTCCTCAACCAACCAACATTCCTCAACCAACCAAC 746
 167 MetProProProProSerGlyProAlaAlaSerSerSerGlyValAlaPro 184
 747 GGAGCGACCGCGCGGAGCTGAGGCAATCGGATGAGTCAAGCGCGCGCG 796
 184 ValAlaSerProGlyProSerGlyAlaSerSerGlyGlyGlyGlyGlyGly 200
 797 CAGGCTCTCTCTGAGTAACTGGCAAGTAAAGTAAAGTAAAGTAAAGTAAAG 846
 201 GlyGlyLeuGlyLeuGlyAlaAlaGlyGlyGlyGlyGlyGlyGlyGlyGly 217

847 GAGGGCAATGCAATGCAAGCTGCTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 896
 217 ValAlaGlyLeuGlySerGlySerGlySerGlySerGlySerGlySerGlySer 244
 897 TGGGGTGGGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 946
 234 AlThrProProProLeuGlyAlaAspProGlyHisSerArgAspAlaArg 250
 947 TTAGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 996
 251 GluArgPheAlaHisPheHisGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 267
 997 CAAGCTTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 1046
 267 ValAlaSerPheAlaGlySerValProGlyPheProGlyHisGlyAlaGly 284
 1047 CTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 1096
 284 SpolInLeuAlaLeuLeuGlyAlaSerHisLeuAlaLeuMetLeuLeuGln 300
 1097 ACCGATTCGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 1146
 301 ThrAlaArgArgTyrAsnHisGlyHisGlyGlyGlyGlyGlyGlyGlyGly 317
 1147 AAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 1196
 317 PheThrTyrSerGlySerPheProHisPheAlaGlyGlyGlyGlyGlyGly 334
 1197 CTGAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 1246
 334 IleLeuAsnProIlePheGlyHisSerArgAlaMetArgArgLeuGlyLeu 350
 1247 TCATCAATTCATCAATTCATCAATTCATCAATTCATCAATTCATCAATTCAT 1296
 351 AspAspAlaGlyGlyTyrAlaLeuLeuValAlaGlyAlaGlyHisSerVal 367
 1297 GAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAG 1346
 367 PheProAsnValGlnGlnProSerArgValGlnAlaLeuGlnGlnPro 384
 1347 GGGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 1396
 384 TyrValGlnAlaArgLeuSerTyrTyrArgTyrGlySerArgProGlnAsp 400
 1397 AATGGAGAGGCTGCTCTCTCAACCAACCAACCAACCAACCAACCAACCA 1446
 401 LeuArgProProArgMetLeuMetLeuGlySerValSerLeuArgHisLeu 417
 1447 CTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 1496
 417 PheValHisSerGlnGlnValPheAlaLeuAlaGlnAspGlySer 434
 1497 CTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 1546
 434 GlnProProLeuLeuSerGlnHisTyrPheValHisGln 446
 1547 TGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 1596
 seq_name: gb_F02.RN014533
 seq_documentation_block:
 1959 1P M882 500 08-NV-1-64
 DEFINITION Ratius norvegicus offhuar receptor mRNA, complete cds.
 ACCESSION 014533
 VERSION 014533.1 GI:565661
 KEYWORDS
 SOURCE Norway rat.
 ORGANISM Ratius norvegicus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Ratius.
 REFERENCE 1 (bases 1 to 1959)

[illegible]

```

      |||
1595 CCTCTGCTGTCCAGATCTGGACGTCACCG 1627
seq_name: qb_pat2:136667

```

```
seq_documentation_block:
```

LOCUS	136667	2000 bp	DNA
DEFINITION	Sequence 1 from patent US 5607967.		
ACCESSION	136667		

KEYWORDS	
VERSION	136667.1
ACCESSION	GI:2086492
130067	

KEYWORDS	:
SOURCE	Unknown.
ORGANISM	Unknown.

UNCLASSIFIED.

REFERENCES
AUTHORS
Friedman, E., Holloway
Vogel, P. I.

TITLE

Treatment of al
carboxylic acid

JOURNAL. Patent: US 5607967-A 1 (14-MAR-1997);
FEATURES Location/Qualifiers

PLANT	LOCATION/CONTAINER
SOURCE	1. , 2030
	for "unbanned"

BASH: COUNT	428 a	655 c	614 g	833
ORIGIN				

```
alignment_scores:
  0.111... 2033.50
  {length}
```

Ratio:	4.751	Gaps
Percent Similarity:	92.842	Percent Ident: 100

alignment block:

US-09-365-576-3 X I 36667 . .

Align seq 1/1 to: 136667 from: 1 to: 2030

```
1 MetSerPro..ThrSerSerLeuAspThrProVal
| | | | | | | | | | | | | | | | | | |
```

245 ATGTCCTCTCGTACACGAGITCGTTGATACACCCCT

```

16 yserProGInProSerThrSerAlaThrSerProThr
| ||||| |::: |::: ||||| |:::

```

295 CCCCCCAGCCCGGGGGCTTCTCTTCAGTCAGCTC

```

33 1yGlnGlu.....ThrasProPro...
   ||  |||
   |||||

```

345 GTCCGAGCCGTTGGCCCGGGGTCCGGACCCTGATGTG

```

44 GlySerSerAlaTyrIleVal.....ValIleGlu
      ::::| | | | | | | | :::: | | | | |

```

395 GAGGCGACGCTCAGCGTGCATTCACATATCGGTCATCCG

```
58 pgluproc[uartq,ysarq,ylslysc,lyproal,aproc]
:|||||
```

445 CGAACCAGTAGCCCAAGCGAAGAGAAAGGTCAGTCTGGA

75 ISGLULECYSARGVALCYSGLYPSPYSALESTGDI
|||||

495 A:GACCTTGGCGTCTCTGTGGACAAAGGCTTCTGG

92 ValLeusctgysrnclycyslsgcylphepricard
|||||

545 GUGU(CAG)UGUGAAGGCT(G) AAGGGCTTCCTTCGGGCT

[illegible]

595 TGGTGGGCTCAAGGCATACCTGGTGGTGGTGGGAA
105 TTATCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG

L25 spaldriemetaiqaiqlysygsimicwysatqike
| | | | | | | | | | | | | | | | | |
6A5 accgttttatccgcacggcaactgaccacctaaccgcctc

042 AUG 11 10AM 1966 SCHEMATA AND ANALOGUES

[illegible]


```

seq_name: qb_p19:HSU14534
seq_documentation_block: 1746 bp 21RA PRI 02-JAN-1995
DEFINITION Human orphan receptor mRNA, partial cds.
ACCESSION U14534
VERSION 014534.1 GI:608134
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens.
Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.
REFERENCE
1 (bases 1 to 1746)
Suzuki, K., Kanki, J.M., Hibiya, F.A. and Iino, S.
Orphan receptor: A novel receptor that modulates gene
expression by retinoic acid and thyroid hormone receptors
Proc. Natl. Acad. Sci. U.S.A. 91 (1994) In press
2 (bases 1 to 1746)
Iino, S.
Direct submission
Submitted (08-SEP-1994) Shun-ichi Iino, from May Institute,
University of Chicago, 5411 S. Maryland, Chicago, IL 60637, USA
FEATURES
location/qualifiers
1..1746
/organism="Homo sapiens"
/sex="male"
/sex="male"
/cell_line="p19-1 prostate carcinoma"
1..1368
/codon_start 1
/product="orphan receptor"
/protein_id "AA05654.1"
/db_xref "GI:608134"
/db_xref "U14534"
FEATURES
polyA_site
746
/feature="eq. A residues"
299 t
BASE COUNT 349 a 571 c 527 g 299 t
ORIGIN
alignment_scores:
Quality: 2019.00 Length: 455
Ratio: 4.773 Gaps: 4
Percent Similarity: 92.967 Percent Identicality: 87.253
alignment_block:
US-09-365-576-3 x HSU14534 ..
Align seq 1/1 to: HSU14534 from: 1 to: 1746
6 SetSeqAlignAP1ProValProGlyAsnGlycSerProGlnProGSerTh 22
|||||
1 ACTTCTCGCAGACAGCCCTGCTGGAAATGTCGACGCTGACCTGACCTG 50
|||||
22 TCGATATGTCCTTThLLeyscLactactylactlu..... 35
|||||
51 CCGCTCTTCTTACCCACGTGAATGAGCAGAGGTCGAGGCGGCGCG 100
|||||
46 .....ThAspProPro...ProGlySerLalysSerSerLalalylr 49
|||||
101 GAGTGGGAGATCTTGATGTCCGACAGTACATTAAGTCAGTCAAGCTG 150
|||||
50 [LEVA].....ValLLeuGluProGlyAspLProGAlaArgLysAr 64
|||||

```

```

151 AACAAAGAGCTGGCTGATCTGCAATATTCGCAAGACCAAGCAAGGCAAGAG 200
64 qlysllysylglytvalprolysmetleucylileucileucylsarvalc 81
|||||
201 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 250
81 ysclyasgllysalaasglleucylleucylleucylleucylleucylleucyl 97
|||||
251 GCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
98 cystylglytylphoaryserysvalvalleucylvalleucylvalleucyl 114
|||||
401 TCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 350
114 valcysvalcyseryseryseryseryseryseryseryseryseryserys 131
|||||
451 TCGCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 400
131 yscyllysleucylleucylleucylleucylleucylleucylleucylleucyl 147
|||||
401 ATTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 450
148 glucysvalleucylleucylleucylleucylleucylleucylleucylleucyl 164
|||||
451 TATTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 500
164 metleucylleucylleucylleucylleucylleucylleucylleucylleucyl 175
|||||
501 GCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 550
175 valalaaseryseryseryseryseryseryseryseryseryseryserys 191
|||||
551 GCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
192 alaaseryseryseryseryseryseryseryseryseryseryseryserys 208
|||||
601 GCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 650
208 acylleucylleucylleucylleucylleucylleucylleucylleucylleucyl 225
|||||
651 TCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 700
225 ysaaseryseryseryseryseryseryseryseryseryseryseryserys 241
|||||
701 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 750
242 aseryseryseryseryseryseryseryseryseryseryseryseryserys 258
|||||
751 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 800
258 valalaaseryseryseryseryseryseryseryseryseryseryseryserys 275
|||||
801 GCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 850
275 valcysvalleucylleucylleucylleucylleucylleucylleucylleucyl 291
|||||
851 GCGGTTCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
292 serthrleucylleucylleucylleucylleucylleucylleucylleucylleucyl 308
|||||
901 TCGATATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 950
308 valcysvalleucylleucylleucylleucylleucylleucylleucylleucyl 325
|||||
951 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1000
325 valcysvalleucylleucylleucylleucylleucylleucylleucylleucyl 341
|||||
1001 TCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1050
342 serthrleucylleucylleucylleucylleucylleucylleucylleucylleucyl 358
|||||
1051 TCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1100

```

```

358 attleallleasullleueseraleasparqproasvalleucylleucyl 375
|||||
1101 GATGGGATCAACATCTCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1150
375 efaarvalleucylleucylleucylleucylleucylleucylleucylleucyl 391
|||||
1151 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
392 thraralleucylleucylleucylleucylleucylleucylleucylleucylleucyl 408
|||||
1201 AGGAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1250
408 lysesvalleucylleucylleucylleucylleucylleucylleucylleucylleucyl 425
|||||
1251 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1300
425 healaleucylleucylleucylleucylleucylleucylleucylleucylleucyl 441
|||||
1301 TCGGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1350
442 trpaspvalleucylleucylleucylleucylleucylleucylleucylleucylleucyl 458
|||||
1351 TGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1400

```

seq_name: qb_hql6:AC073806

seq_documentation_block:

LOCUS AC073806 215601 bp DNA HTG 18-JUN-2000
 DEFINITION Mus musculus clone Rp23-47N6, WORKING DRAFT SEQUENCE, 22 ordered
 pieces.

ACCESSION AC073806

VERSION AC073806.2 GI:9256408

KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus; Mus-

REFERENCE JDF Joint Genome Institute.

TITLE Sequencing of Mouse

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 215601)

AUTHORS JDF Joint Genome Institute.

TITLE Direct Submission

JOURNAL Submitted (29-JUN-2000) ProteinSeq Sequencing Facility, JDF Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 on Jul 18, 2000 (this sequence version replaced at:881842).

-----Genome Center

Center: Joint Genome Institute

Center Code: JGI

Web site: <http://www.jgi.doe.gov>

Project Information

Center Project Name: 175238

Center clone name: Rp21-23_47N6

Summary Statistics

Consensus quality: 203615 bases at least Q40

Consensus quality: 211090 bases at least Q40

Consensus quality: 212563 bases at least Q40

Estimated insert size: 220000; agarose-gel estimation

Estimated insert size: 214601; sum-of-counts estimation

quality coverage: 7.88 in Q20 bases; agarose-gel estimation

quality coverage: 8.1 in Q20 bases; sum-of-counts estimation.

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 22 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and

```

* the accession number will be preserved.
1 4864: contig of 4864 bp in length
* 4865 4864: gap of unknown length
* 21161 21161: contig of 21161 bp in length
* 21162 21161: gap of unknown length
* 21163 21162: contig of 812 bp in length
* 22074 22173: gap of unknown length
* 22174 33484: contig of 11311 bp in length
* 33485 33585: gap of unknown length
* 33586 33628: contig of 5944 bp in length
* 33629 33629: gap of unknown length
* 33630 62157: contig of 22529 bp in length
* 62158 62157: gap of unknown length
* 62159 76745: contig of 14488 bp in length
* 76746 76845: gap of unknown length
* 76846 81102: contig of 4257 bp in length
* 81103 81202: gap of unknown length
* 81203 85706: contig of 4501 bp in length
* 85707 85806: gap of unknown length
* 85807 90755: contig of 4949 bp in length
* 90756 90855: gap of unknown length
* 90856 101543: contig of 10688 bp in length
* 101544 119328: contig of 17685 bp in length
* 119329 119426: gap of unknown length
* 119427 122789: contig of 3362 bp in length
* 122790 131818: contig of 9029 bp in length
* 131819 131918: gap of unknown length
* 131919 150107: contig of 18189 bp in length
* 150108 150207: gap of unknown length
* 150208 156468: contig of 6261 bp in length
* 156469 156969: gap of unknown length
* 156970 187823: contig of 30755 bp in length
* 187824 187824: gap of unknown length
* 187825 190416: contig of 2603 bp in length
* 190417 190516: gap of unknown length
* 190517 195937: contig of 5420 bp in length
* 195938 196037: gap of unknown length
* 196038 204419: contig of 8382 bp in length
* 204420 204518: gap of unknown length
* 204519 213898: contig of 9480 bp in length
* 213899 213998: gap of unknown length
* 213999 215601: contig of 1603 bp in length
FEATURES
SOURCE
1. 215601
Location/Qualifiers
/organism="Mus musculus"
/db_xref="taxon:10090"
/feature="K023-47N6"
/feature="K023-47N6"
/feature="K023-47N6"
BASE COUNT 52000 a 52880 c 58600 g 50020 t 2101 others
ORIGIN
alignment_scores
Quality: 1929.00 Length: 918
Ratio: 4.905 Gaps: 11
Percent Similarity: 48.584 Percent Identity: 48.136
alignment_block:
US-09-365-576-3 X AC073806
Align Seq 1/1 to: AC073806 from: 1 to: 215601
1 MotSeqSerProthSerSerLeuAspThrProVal..... 12
|||||
111862 ATGCTCTCCGCAACAGCTCTGCAACTCCCTCCGCGAGAGAG 111911
12 ..... 12
111912 GCGTTGCTAGCCAGCCCTTCACAGAGTGAGAGAACTACTGCTGCT 111961
13 .....ProGlyncylSerProGlnProSerThrSerAlaIsh 25

```

```

112862 GCGAAGCAATAGCCAGGAGCTGGAGGAAAGAGAGGATGACATGAC 112911
207 AATAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 223
112912 GCGCTGAGAGTGTATGATGATGATGATGATGATGATGATGATGATG 112961
223 SASNLYSATGSGTPhosphoAspGluProLysValThr..... 235
112962 CAGGCAAGCAATTTTCTGATGATGAGCCAAATGATGGATGACCTGCG 113011
235 ..... 235
113012 CAGCTCTGTCACACATCCCATGAGCTTAAAGAGAGAGGTTCTTCA 113061
236 ..... 236
113062 TCGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 113111
239 ..... 239
113112 GCGTGAAGAGTGTATGATGATGATGATGATGATGATGATGATGATG 256
113112 GCGTGAAGAGTGTATGATGATGATGATGATGATGATGATGATGATG 113161
256 h-PhosphoValAlaThrLeuValGlnGluGluValAspPheAlaLys 272
113162 TCAAGAGCTTACGACATGATGATGATGATGATGATGATGATGATGATG 113211
273 GluValPheGluPheGluPheGluPheGluPheGluPheGluPheGlu 289
113212 CAGGCTGAGAGTGTCTCTGATGATGATGATGATGATGATGATGATGAT 113261
289 uLysAlaSerThrThr..... 294
113262 GAGGAGGTCATGATGATGATGATGATGATGATGATGATGATGATG 113311
294 ..... 294
113312 GCGAAGCAATAGCCAGGAGCTGGAGGAGCTGGAGGAGCTGGAGCTGG 113361
294 ..... 294
113362 AGGCTTCATGATGATGATGATGATGATGATGATGATGATGATGATG 113411
294 ..... 294
114412 GAAAGAGCAATAGGCTTACGATGATGATGATGATGATGATGATGATG 114461
294 ..... 294
114462 TGTAAAGAGCAATAGCTTAAAGAGCTTAAAGAGCTTAAAGAGCTT 114511
294 ..... 294
114512 CAAATGAGATGATGATGATGATGATGATGATGATGATGATGATG 114561
294 ..... 294
114562 AAGCTTAAGATTAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTT 114611
294 ..... 294
114612 TCGATTTGATGAGGAGAGTGGATGATGATGATGATGATGATGATG 114661
294 ..... 294
114662 AAGCTTAACTGAGCAAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 114711
295 ..... 295
114712 ATCTGGAGTAACTGATGATGATGATGATGATGATGATGATGATGATG 114761
298 ..... 298
114762 GCTAAGAAAGAGTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCT 114811

```

```

315 cULysAspPheThrTyrSerLysAspPheHisValAlaCys.....Leu. 430
113812 TAAAGGATTCATGATGATGATGATGATGATGATGATGATGATGATG 113861
330 ..... 430
113862 AAGGAGTACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 113911
330 ..... 430
113912 GCGAGGAGCTGCGAAGAAATCATGTTAAAGAGAGAGAGATGATGAT 113961
330 ..... 430
113962 GTAAGAGTACTACAGAGTCTCTGTTGAGAGGAGAGAGTGAAGTAC 114011
330 ..... 430
114012 ACAGGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 114061
331 ..... 430
114062 TGAAGCTGAGCTGAGAGGAGCTCTCTCTCTCTCTCTCTCTCTCTCT 114111
336 ..... 430
114112 TCCATCTTGGAGTGTCTGAGAGGAGATGAGAGGAGTGGAGTGGAGCA 114161
353 h-AspLysAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 469
114162 CAGAGTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 114211
370 AsnValGlnGluProSerArgValGluAlaLeuGluGluGluProTyrVal 486
114212 AATGCGAGAGAGTCAAGCGGTGTGAGAGGAGTCAAGAGTCAAGAG 114261
386 uAlaLeuLeuSerTyrThrArgGluLysArgPro..... 497
114262 GCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 114311
397 ..... 497
114312 GCTGTATATGATGATGATGATGATGATGATGATGATGATGATGATG 114361
397 ..... 497
114362 CACACTGAGGCTGAGAGGCTGAGGAGCTTAAAGAGTGTGGCTTTC 114411
397 ..... 497
114412 GCTGCTTTGCAAGACAGAGGAGTGAAGAGAGAGAGAGAGAGAGAG 114461
398 ..... 497
114462 TCGAGAGCAGGCTGCGCTTTCAGGATGATGATGATGATGATGATGATG 114511
414 ArgThrLeuSerSerValHisSerGluGluGluGluGluGluGluGlu 430
114512 GCGAGCTGAGCTGCGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 114561
430 uAspLysLysLeuProProLeuLeuLeuSerGluLeuLeuPheAspVal 446
114562 GAGCAAGAGAGTGTGAGCTTGTCTGTGATGATGATGATGATGATG 114610
seq_name: qb_r1:MM0132602
seq_documentation_block:
locus: MM0132602 6865 bp DNA Ref 03-MAR-2000
DEFINITION Mus musculus LXR-beta gene, exons 1-10.
ACCESSION AJ132602
VERSION AJ132602.1 GI:5689852
KEYWORDS LXR beta gene; nuclear oxysterol receptor; LXR-beta.

```



```

|||||
4533 ATGTCCTGGCCCAAGATCTCTGAGCACTGCGGTGGCTGGTGAATTACG 4582
12 .....
4584 GAGCTTTCCTAACCAATCCCTCTCAATAGCT GAGCAAGCTCACTGGCT 4642
13 ..... Pro-Gly-Ser-Gly-Ser-Pro-Gly-Pro-Ser-Thr-Ser-Ala-Thr 25
4643 GTCTTCCTTTCTCTAGGCAATGGTTCTCTCAATCCAGTACCTCTGGTACG 4682
25 rSer-Pro-Thr-Ileu-Ser-Ileu-Gly-Ileu-Gly-Ileu-Thr-Asp-Pro-Thr-Pro-Lys 42
4684 GTCAATCACTAATTAAGCAAGAGGAGGAGAGAGAGAGAGAGAGAGAGAG 3742
42 GTCTTCCTTTCTCTAGGCAATGGTTCTCTCAATAGCTGAGATGGGAGAGAGAA 3782
50 .....
4784 GAGGCTCTGGAGCTTTCTTGAGAGAGAGGAGAGAGAGAGAGAGAGAGAG 3842
50 .....
4884 GTTTCCTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3882
50 .....
4884 GCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3942
50 .....
4944 TCTGAGCTGAGCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3982
51 ..... Val...Val-Ileu-Glu 55
4984 CTCTTCTGCTTAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4042
55 GPro-Gly-Asp-Gly-Pro-Gly-Ileu-Gly-Ser-Ileu-Gly-Ser-Ileu-Gly-Ser 72
4044 GGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4082
72 CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 88
4084 TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4142
89 HSTPTASAVALLHNSGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 105
4144 CACTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4182
105 TAAVALLHNSCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 122
4184 TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4242
122 yscIme1AspAlaPheMetArgArgLysCysIleuLeuGlyArgLysArg 148
4244 GGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4282
149 LysCysLysCysLysLysLysLysLysLysLysLysLysLysLysLysLys 150
4284 AAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4342
150 .....
4444 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4482
151 ..... LeuSerGluGluGluIleuArg 157
4484 GAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4542
157 GlysLysArgLysLysLysLysLysLysLysLysLysLysLysLysLysLys 174
|||||

```

```

4443 GAAGAAAGAGATTCAAGAACAGCAACAGAGAGAGAGAGAGAGAGAGAGAG 4482
174 LProAlaAlaSerSerSerGlyArgProAlaAlaSerProGlyThrSer 190
4483 AGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4542
191 GluAlaSerSerGluLysSerGlyLysLysLysLysLysLysLysLysLys 207
4543 GAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4582
207 AlaSerGluLeuMetIleGluIleuLeuValAlaAlaIleuLeuThrGly 224
4583 GGTCTAGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 4642
224 GlysArgSerPheSerAspLysValIlePhe..... 245
4643 ACNAAAGATTTTCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4682
235 .....
4683 GCGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4742
236 ..... ProIleProLeuG 240
4743 GAGATCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4782
240 LValAspProGlnSerArgAspAlaArgGluArgGlnArgPheAlaHisPhe 256
4783 GTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4842
257 ThrGluLeuAlaIleIleSerValGluGluIleValAspPheAlaLysGlu 274
4843 ACGAGCTAGGCTATCATCTGGGTCTGAGAGAGAGAGAGAGAGAGAGAGAG 4882
273 PValIleProGlyPheLeuGlnIleuGlyArgGluAspGlnIleuAlaLeuLeu 290
4884 GGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4942
290 ysaIAscThrIle..... 294
4943 AGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4982
294 .....
4983 GAACATGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5042
294 .....
5043 GGTCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5082
294 .....
5083 AAAGGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5142
294 .....
5143 TTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5182
294 .....
5183 AATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5242
294 .....
5243 GGTAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5282
294 .....
5283 GGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5342
294 .....
5343 GTCTTACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5382

```


cc plasmid pGEN-2, as an EcoRI fragment, to form pGEN-1.5p6. The
cc encoded protein is involved in modulation of retinoic acid
cc signaling.

xx Sequence 1944 bp: 440 A: 588 G: 544 G: 362 T: 0 other:

alignment_scores:

quality: 224.00 length: 416
ratio: 5.121 gaps: 0
Percent Similarity: 100.000 Percent Identity: 98.286

alignment_block:

US-09-365-576-3 x T13229

Align seq 1/1 to: 113229 from: 1 to: 1934

```

1 MetSerSerProThrSerSerLeuAspThrProValProAlaProAlaLysLeu 17
|||||
191 ATGCTTCTCCCAATAGCTCTGGACATCTGCTTGGCTTGGCAATGGATGTTG 240
17 ProGlnProSerThrSerAlaThrSerProThrIleLysGlnGlnGly 34
|||||
241 TCCCAAGCCAGTACGATCTGCTGACCTTCACCTATTAAAGAGAGAGAGAG 290
34 LysGlnThrAspProProGlnLysSerGlnLysSerSerAlaThrIle 50
|||||
291 AGCAATCTGATCCATCCACAGCTCTGACAGCCCTGACAGCTCTGCTTACATC 400
51 ValValIleLeuGlnThrProLysAspGlnProGlnThrAlaThrLys 67
|||||
341 GTGCTGATCTTAGAGATAGAGAGATGAGATGAGATGAGATGAGATGAGAT 450
67 TyrAlaAlaProLysMetLeuGlnLysIleLysGlnLysAlaThrVal 84
|||||
391 TCCGCGCCGCAAGATGCTGAGGATGAGATGCTGCGCGCTGCGGAGACA 440
84 LysAlaLysSerGlnThrProLysAspValLysSerGlnGlnGlnGlnGln 100
|||||
441 AAGCTTCTCCCAATAGCTCTGGACATCTGCTTGGCTTGGCAATGGATGTTG 490
101 PhePheAlaThrSerProThrSerAlaThrSerProThrIleLysGlnGln 117
|||||
491 TTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
117 GcLysSerGlnThrProLysMetAspAlaThrMetAlaThrAlaLysSer 144
|||||
541 GGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 590
134 GlnYsaAlaGlnLysSerGlnLysGlnAlaThrMetAlaThrGlnLysVal 150
|||||
591 TCTGATACCTGCTGCAATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 640
151 LeuSerGlnGlnGlnIleAlaThrLysAlaThrIleGlnLysGlnGlnGln 167
|||||
641 CTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 690
167 GcLysSerGlnThrProLysMetAspAlaThrMetAlaThrAlaLysSer 184
|||||
691 GAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 740
194 LysAlaThrProGlnThrSerGlnAlaLysSerGlnIleLysSerGlnGln 209
|||||
741 CAGCTTCTCCCAATAGCTCTGGACATCTGCTTGGCTTGGCAATGGATGTTG 790
201 GlnGlnLysIleGlnLysAlaThrAlaThrAlaThrIleGlnLysGlnGln 217
|||||
791 GAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
217 LysAlaThrGlnLysGlnLysAlaThrSerPheSerAspThrProLysVal 244
|||||
841 TCTGATACCTGCTGCAATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 890

```

```

244 ATThrProThrProLeuAlaAspProGlnLysSerAlaThrAlaThrGln 250
|||||
891 TCCGCGCCGCAAGATGCTGAGGATGAGATGCTGCGCGCTGCGGAGACA 940
251 GlnAlaThrPheAlaThrSerProThrSerAlaThrSerProThrIleLys 267
|||||
941 CAGCTTCTCCCAATAGCTCTGGACATCTGCTTGGCTTGGCAATGGATGTTG 990
267 ValAspPheAlaThrSerProThrSerAlaThrSerProThrIleLysGln 284
|||||
991 CTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1040
284 SpGlnIleAlaThrProLysAlaThrIleGlnLysMetLeuLysGln 400
|||||
1041 AGCAATCTGATCCATCCACAGCTCTGACAGCCCTGACAGCTCTGCTTACATC 1090
301 ThrAlaThrAlaThrSerProThrSerAlaThrSerProThrIleLysGln 417
|||||
1091 ACAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
317 PheThrTyrSerLysAspPhePheAlaThrAlaThrAlaThrValGln 344
|||||
1141 CTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1190
344 IleLeuAspProThrProGlnLysSerAlaThrMetAlaThrAlaThrGln 450
|||||
1191 TCAATCAATCCATCTGAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1240
351 AspAspAlaThrProLysAlaThrAlaThrAlaThrAlaThrAlaThrAla 467
|||||
1241 GAGCAATGCTGCAATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1290
467 PheProAspValIleGlnLysSerAlaThrAlaThrAlaThrAlaThrAla 484
|||||
1291 TCCGCGCCGCAAGATGCTGAGGATGAGATGCTGCGCGCTGCGGAGACA 1340
384 TyrValGlnAlaThrProLysSerGlnThrAlaThrLysAlaThrProGln 400
|||||
1341 ATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1390
401 LeuAlaThrProAlaThrMetLysAlaThrAlaThrAlaThrAlaThrAla 417
|||||
1391 CTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
417 SerValIleSerGlnIleValPheAlaThrAlaThrAlaThrAlaThrAla 444
|||||
1441 CTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1490
444 GlnProLeuLeuSerGlnIleThrAspValIleGln 446
|||||
1491 TCTGATACCTGCTGCAATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1528

```

seq_name: US-09-365-576-3
seq_documentation_block:
ID: Q88761 standard: DNA: 1959 bp.
AC: Q88761:
XX
AC: Q88761:
XX
DT: 20-JAN-1996 (first entry)
DE: Rat ubiquitous nuclear receptor protein RNA.
KW: Ubiquitous nuclear receptor drug dosing; disease diagnosis;
KW: therapy; ss.
XX
OS: Rattus rattus.
XX
PH: Key location/Qualifiers
FT: 256..1919
FI: /tag a
XX
FN: W09513373-A1.

XX 18 MAY-1995.
 XX 08 NOV-1994: 94MO-US19884
 XX 10 NOV-1994: 9408-0152004
 XX (ARCH.) ARCH DEV CORP.
 XX Liao S. Song C.
 XX WPI: 1995, 194092/25.
 XX P-PSDB: R74749.
 XX New ubiquitous oncoprotein receptor: used to develop precls. for use in
 XX diagnosis, drug design and therapeutic applications
 XX Claim 6: Page 151: 199pp: English.

This genomic DNA isolated from rat cDNA may be cloned
 recombinantly to produce a cDNA which interacts with the response
 elements and network of receptors in the thyroid hormone/protein
 and receptor subfamily. The products can be used in detection,
 diagnosis and in screening assays for substances which interact
 with cDNAs for use in diagnosis, drug design and therapeutic
 applications.

Sequence 1959 bp: 406 A, 607 C, 565 G, 381 T, 0 other.

alignment_scores:
 quality: 2247.50 length: 446
 Ratio: 5.085 gaps: 1
 Percent Similarity: 99.104 Percent Identity: 97.085

alignment_block:
 US-09-365-576-3 x Q88761 ..

Align seq 1/1 to: Q88761 from: 1 to: 1959

1 MetSerSerProThrSerSerLeuAspThrProValProGlyAsnGlySer 17
 256 ATGCTCTTCCCAAGATGCTCTGACACTGCTTCTCCGGAATGGTGC 305
 17 PProGluProSerThrSerAlaThrSerProThrTleuGlyGluGlyGly 34
 406 TCTGCAAGCTCAGTACGCTGCTGCACTTCAACCACTATTAAGGAGGAGTAC 355
 44 InclThrAspProProGluGlySerGlyGlySerSerSerAlaTyrIle 50
 456 AAGACATATATCACTTCACCTTCGAGGCTTCACCTGCTTACATC 405
 51 ValValIleIleLeuGluProGluAspGluProGluIleValAsnGlySer 67
 406 GTG.....GAGGCAAGACATGACCTGACGCAAGGCAAGACAGG 446
 67 yProAlaProGlyMetIleuIleIleIleGlySerAspValGlySerAsp 84
 447 TCGGAGGCAAGATGCTGAGTATGAGCTCTGCAAGCTTGGCGGAGAA 496
 84 ysaIleSerGlyPheGluIleTyrAsnValLeuSerGlyGluGlyGlySer 100
 497 AAGCTCTGAGCTTCCAGTAAATAGTCTCAGCTGCAAGGCTGCAAGAG 546
 101 PhePheArgArgSerValValIleGluGlyIleGluGluTyrAlaGlySer 117
 547 TTCTTTCGAGGCTAGGCTGCTGCTCATGCTGAGCGGAGGCTTATGCTG 596
 117 gglYserGlyThrGlySerIleMetAspAlaPheMetArgAspGlyGlyGlu 134
 597 GGGCAAGGAGAGCTGCTGAGCATGATGCTTCATAGGCGGAGTGGCAGC 646
 134 GngYsaIleGluArgTyrGlySerIleValGluIleMetArgIleGluGlyVal 150

151 IouSerGluGluGluIleIleArgTyrIleAspIleGluGluGluGluGlu 167
 647 TGTGACAGCTGGGCAAGTGTAAAGAGGCTGATGAGGAGAGTGGTGG 696
 647 CTTCCTGAGCAATCAATGCAAGAAAAAATTCAGAAAGCAATCAATCA 746
 167 ngIleProProProSerGluProAlaIleSerSerSerGlyIleArgAs 184
 747 GCAAGCAGCGCGCGGAGCTGAGCAGCAGCATGTGGTACGTGAGGCTG 796
 184 IAlaSerProGlyThrSerAlaIleSerSerGlyGlySerIleGlyGly 200
 797 CAGCTGCGGCTGGCACTTGGAGAGCAAGTATCAAGGCTGCGGAGAA 846
 201 GluGlyIleGluIleuThrAlaIleGluGluMetIleGluGluGlyVal 217
 847 GAGGAGATCCAGCTGACAGCGGCTGCAAGCTGATGATGATCAATGAT 896
 217 IAlaIleGluIleGluGlySerValIleSerSerSerSerSerSerSer 234
 897 TCCGCTGCACTGCTGATGATCAAGGAGATCTTTCGAGCAGCTTAAG 946
 234 AlThrProIleProGluGlyAlaAspProGluIleSerArgAspAlaArg 250
 947 TCAAGGCTGCGCTTGGCTGAGAGCTGAGTGGGAGCAAGCTGGTAG 996
 251 GluArgPheAlaHisPheThrGluGluAlaIleIleSerValGluGlu 267
 997 CAAGGCTTGGCTATTTAGTGAAGTATATATATATATATATATATAT 1046
 267 GValAspPheAlaIleGluIleValIleProGlyPheGluGluGluGlu 284
 1047 GAGGAGCTTGGCTGAGTGGTGGAGAGTGGTGGAGTGGTGGAGTGG 1096
 284 spGluIleAlaLeuLeuGlyAlaSerThrIleGluIleMetLeuLeuGlu 300
 1097 ACCAGATGCGCTCTGTAAGAGCATGCAATGAGATGATGCTTGTAG 1146
 301 ThrAlaArgArgTyrAsnHisGluThrGluGlySerIleThrPheGluGly 317
 1147 AAGGCAAGAGCTGACAGCAAGCAAGAGTATGATGATGATGATGAT 1196
 317 PheThrTyrSerGlyAspAspPheHisArgAlaGluGluValGluGlu 134
 1197 CTTCAGCTAGCAGCAAGCAGCAGCTTGCAGCTGAGGCTTGCAGCTG 1246
 334 IleIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGluGlu 350
 1247 TATCATATCCCATCTTATGATTCTCTGAGGCTATGAGTGGGTA 1296
 351 AspAspAlaGluIleValAlaLeuLeuIleAlaIleAsnIlePheSerAla 367
 1297 GAGCAGTACAGAGTATGCTTGTCTGCTGATGCAATGCAATTTGTAG 1346
 367 ParqProAsnValGluGluProSerArgValIleValAlaLeuGluGlu 384
 1347 CCGAGCTTAAAGTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1396
 384 yValGluAlaLeuLeuSerTyrThrArgIleIleGlySerProGluAsp 400
 1397 ATGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1446
 401 IouArgPheProArgMetLeuMetLeuValSerLeuValSerLeuVal 417
 1447 CTGCGCTTGGCAGAGATGCTCATGAAAGTGTGAGGCTGGGATCTG 1496
 417 rSerValHisSerGluGluValPheAlaLeuArgLeuGluAspGlySer 434
 1497 CTCTGTGCACTGCGAGCAGGCTTTCGATATGCTGTTCAGAGCAAT 1546
 434 euProProLeuLeuSerGluIleTyrAsnValHisGlu 446

Quality: 2034.50 Length: 461
 Ratio: 4.751 Gaps: 5
 Percent Similarity: 92.842 Percent Identity: 87.202

alignment block:
 US-09-365-576-3 x T30041 ..

Align seq 1/1 to: T30041 from: 1 to: 2030

```

1 MetSerPro...ThrSerLeuAspThrProValProGlyAsnGlu 16
  |||||
245 ATCTCTCTCTACGACGAGTTCCTGGTACACCGCTGCTGGTGGAAATGG 294

16 ySerProGluProSerThrSerAlaThrSerProIleHisGluGln 33
  |||||
295 TCTCTCTACGACGAGTTCCTGGTACACCGCTGCTGGTGGAAATGG 344

33 yGluGln.....ThrAspProPro...ProGlySerGlu 43
  |||||
445 GTCTGGACGAGTTCCTGGTACACCGCTGCTGGTGGAAATGG 394

44 GlySerSerAlaIleVal.....ValIleGluIleProGluAs 58
  |||||
495 GAGGCAATCTCAGCTTCAGACACACCTGATTCATATTCCTGACAA 444

58 pGluProGluArgLysArgLysGlyProAlaProGlySerAlaGly 75
  |||||
445 GAGGCAATCTCAGCTTCAGACACACCTGATTCATATTCCTGACAA 494

75 SerLeuGlySerValGlySerValGlySerValGlySerValGly 91
  |||||
495 AAGAGCTTTCCTGCTGCTGGTACACCGCTGCTGGTGGAAATGG 544

92 ValLeuSerGlyGluGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 108
  |||||
545 GTCTCTCTCTACGACGAGTTCCTGGTACACCGCTGCTGGTGGAA 594

108 pGlyGlyValGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 125
  |||||
595 TCTCTCTCTACGACGAGTTCCTGGTACACCGCTGCTGGTGGAA 644

125 pAlaPheMetArgLeuGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 141
  |||||
645 AAGCTTCTCTCTACGACGAGTTCCTGGTACACCGCTGCTGGTGG 694

142 GluAlaGluMetArgLeuGlyGlyGlyGlyGlyGlyGlyGlyGly 158
  |||||
695 GAGGCAATCTCAGCTTCAGACACACCTGATTCATATTCCTGAC 744

158 slySerGluGluGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 169
  |||||
745 GAGGCAATCTCAGCTTCAGACACACCTGATTCATATTCCTGAC 794

169 rGluProProSerGluProAlaAlaSerSerSerGlyArgProAla 185
  |||||
795 GAGGCAATCTCAGCTTCAGACACACCTGATTCATATTCCTGAC 844

186 SerProGlyThrSerGluAlaSerSerSerGlyGlyGlyGlyGly 202
  |||||
845 TCTCTCTCTACGACGAGTTCCTGGTACACCGCTGCTGGTGGAA 894

202 yLleGluGlnThrAlaAlaGluGlnGlnGlnGlnGlnGlnGlnGln 219
  |||||
895 TCTCTCTCTACGACGAGTTCCTGGTACACCGCTGCTGGTGGAA 944

219 LeuLeuGluGlnGlySerArgSerPheSerArgGluProGlyVal 235
  |||||
945 TCTCTCTCTACGACGAGTTCCTGGTACACCGCTGCTGGTGGAA 994

236 ProThrProGluGlyAlaAspProGluSerArgArgAlaArgGln 252
  |||||
995 GTCTGGACGAGTTCCTGGTACACCGCTGCTGGTGGAAATGG 1044

```

```

252 pPheAlaHisPheThrGluAlaLeuLeuSerValGluGluIleVal 269
  |||||
1045 GTCTGGACGAGTTCCTGGTACACCGCTGCTGGTGGAAATGG 1094

269 pPheAlaHisPheThrGluAlaLeuLeuSerValGluGluIleVal 285
  |||||
1095 GTCTGGACGAGTTCCTGGTACACCGCTGCTGGTGGAAATGG 1144

286 pLeuAlaLeuGlyAlaSerThrIlePheIleMetLeuGluIleAla 302
  |||||
1145 ATCTCTCTCTACGACGAGTTCCTGGTACACCGCTGCTGGTGGAA 1194

202 pArgArgTyrAsnHisGluThrLeuSerIleGluIleLeuGlySer 319
  |||||
1195 GAGGCAATCTCAGCTTCAGACACACCTGATTCATATTCCTGAC 1244

319 ThrTyrSerLysAspAspPheHisArgAlaTyrGluGlnValGlu 345
  |||||
1245 GTCTGGACGAGTTCCTGGTACACCGCTGCTGGTGGAAATGG 1294

336 AsnProIlePheGluPheSerArgAlaMetArgLeuGlyLysAsp 352
  |||||
1295 GAGGCAATCTCAGCTTCAGACACACCTGATTCATATTCCTGAC 1344

452 pAlaGluTyrAlaLeuLeuLeuAlaIleAsnIlePheSerAlaAsp 369
  |||||
1345 GTCTGGACGAGTTCCTGGTACACCGCTGCTGGTGGAAATGG 1394

369 rAsnValGluGlnProSerArgValGluAlaGluGlnGlnProTyr 385
  |||||
1395 GAGGCAATCTCAGCTTCAGACACACCTGATTCATATTCCTGAC 1444

386 GluAlaLeuLeuSerTyrThrArgIleGlyArgProGluAspGlu 402
  |||||
1445 GTCTGGACGAGTTCCTGGTACACCGCTGCTGGTGGAAATGG 1494

402 pPheProArgMetLeuThrLysLeuValSerLeuArgThrGlySer 419
  |||||
1495 GTCTGGACGAGTTCCTGGTACACCGCTGCTGGTGGAAATGG 1544

419 pAluHisGluGlnValPheAlaLeuAlaLeuIleAspLysLysLeu 435
  |||||
1545 GTCTGGACGAGTTCCTGGTACACCGCTGCTGGTGGAAATGG 1594

436 ProLeuLeuSerGluIleThrAspValHisGlu 446
  |||||
1595 GTCTGGACGAGTTCCTGGTACACCGCTGCTGGTGGAAATGG 1644

seq name: US-09-365-576-3 to: seq name: P3-NV-096 CAT T27616
seq description: block:
ID: T27616 standard, cDNA to mRNA, 1579 bp.
XX
AC T27616
XX
DI 09-NV-1996 (first entry)
XX
DE Human foetal lung steroid hormone receptor analogue EC2B cDNA.
XX
KW Human; foetal lung; steroid hormone; receptor; analogue protein;
KW EC2B protein; cDNA; sequence; binding domain; foetal; foetal;
KW foetal; foetal; foetal; foetal; foetal; foetal; foetal; foetal;
KW probe; antibody; immunohistochemical assay; ds.
OS
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS
FT /tag= d
XX
PN W09009424-A1.
XX
PD 28-MAR-1996.

```

```

XX 21-SEP-1995: 95W-1-JP01-009.
XX
XX 21-SEP-1994: 94JH-0226270.
XX
XX (EISA) EISA CO LTD.
XX (GANN) 2H GAN KIMURA.
XX (GANN) CANCER INST.
XX
XX Makumura Y, Saito H:
XX
XX WP1: 1996-188403/19.
XX P-ESDB: K96234.
XX
XX ECDN protein, a steroid hormone receptor analogue from human foetal
XX lung - is expressed in cancer cells and is useful for cancer
XX diagnosis and drug development.
XX
XX claim 1: Pages 21-25; 44pp; Japanese.
XX
XX The present sequence encodes the human foetal lung derived steroid
XX hormone receptor analogue protein, ECDN. In various cancer cells a
XX variant ECDN protein, designated ECDN small mol. (ECNsm) protein,
XX is expressed. Therefore screening for ECDN and ECNsm protein
XX binding moles, using recombinant ECDN and ECNsm proteins will be
XX useful in the identification of candidate anticancer drugs. Gene
XX expression of ECDN and ECNsm proteins in normal and cancerous
XX tissues can be studied using primers and probes derived from ECDN
XX and ECNsm protein cDNA. Antibodies which recognise ECDN and ECNsm
XX immunohistochemical assays.
XX
XX Sequence 1979 bp; 405 A; 642 C; 602 G; 340 T; 0 other;
XX
alignmeot_scores:
XX quality: 2926.50 length: 461
XX Ratio: 4.746 gaps: 5
XX Percent Similarity: 92.625 Percent Identity: 86.985
XX
alignmeot_block.
XX US-09-365-576-3 x 127616
XX
Align seq 1/1 for T27616 from 1 to: 1979
XX
1 Met Ser Ser Pro The Ser Ser Leu Asp Thr Pro Val Pro Gly Asn G1 16
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
206 ATGTTCTTCTCTATCCAGAGTTCCCTGGATAGCCCGCTGCTGGAAATGG 255
16 Y Ser Pro Thr Pro Ser Thr Ser Ala Thr Ser Pro Thr Thr Leu Ser Thr G1 33
1 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
256 CAGCTTCAGAGTGGCGGCTCTCTCTCTACACAGTGAAGGAGGAGGAGG 305
43 Lys Glu Thr.....Thraspropro...Pro Gly Ser Glu 43
1 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
406 GTTGGAAATCTGTAATTCGAGAGCTGATGTCCTGAGGACATGAT 355
44 Gly Ser Ser Ser Ala Thr Thr Val.....Val Thr Leu Thr Pro Glu As 58
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
406 GAGGTCATGTCAGCTTGTAGACACAGTGGGTGATCGACATCGCCGAGAA 405
58 P Glu Thr Pro Thr Thr Thr Ser Thr Thr Thr Thr Thr Thr Thr Thr 75
1 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
406 GCAATCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 455
75 Lys Glu Thr Ser Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr 91
1 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
456 ATGAAATTGCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 505
92 Val Thr Ser Ser Ser Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr 108
1 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
506 GTGTCTGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 555

```

```

108 GGTGCTGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 125
1 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
556 TGGTGGGGGTAGGGGCTATGGCTGGTGGGGGTAGGGGAGGAGGAGGAGGATGG 605
125 SPATATPMeTATGATATGATATGATATGATATGATATGATATGATATGATAT 141
1 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
406 AGCTTTTATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 655
142 GUAAGAGTMeTATGATATGATATGATATGATATGATATGATATGATATGATAT 158
1 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
656 GAGGTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 705
158 STYATATLeTLeTLeTLeTLeTLeTLeTLeTLeTLeTLeTLeTLeTLeTLeTLeT 169
1 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
706 GAAAGAAATTCGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 755
169 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 185
756 GAGCTGTGGGGGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 805
186 STCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 202
806 TCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 845
202 YTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 219
1 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
856 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 905
219 LCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 245
906 GCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 955
236 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 252
956 GCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1005
252 STHeATATLeTLeTLeTLeTLeTLeTLeTLeTLeTLeTLeTLeTLeTLeTLeTLeT 269
1006 CTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1055
269 STGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 285
1056 ATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1105
286 TLeATLeTLeTLeTLeTLeTLeTLeTLeTLeTLeTLeTLeTLeTLeTLeTLeT 302
1106 ATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1155
302 ATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 319
1156 GAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1205
319 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 345
1206 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1255
336 ASGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 352
1256 AACCCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1305
352 TATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 369
1306 GCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1355
369 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 385
1356 GTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1405
386 GTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 402
1406 GAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1455
402 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 419

```

```

|||||
1456 CTTCGGGCGATGCTCATGAGAGCTGTCAGCTGGGCAAGCTTACGCTCTG 1505
419 ahhssscgldhslvalflcalalcalalqlcalclalaspysyslscd 435
|||||
1506 TGACATCGGATCAAGTCTTCGCTTTCAGGCTCCAGACCAAGAAAGCTGGG 1555
446 Proleuleusgclulctfpaspvaliscl 446
|||||
1556 CCTCTGCTGTCAGAGATCTGACGCTGACGAG 1588
seqname: us09-365-576-3 from: 1 to: 1898, 1898, 1898
seq_documentation_block:
ID 179634 standard: DNA: 1898 bp.
XX
AC 179634:
XX
DI 15-061-1997 (first entry)
XX
DE DNA encoding human ubiquitous nuclear receptor polypeptide.
XX
KW Ubiquitous nuclear receptor polypeptide; DNA binding domain;
XX
KW therapy; drug design; diagnosis; ss.
XX
OS Homo sapiens
XX
FH Key Location/Qualifiers
FT CDS 71..1453
ET 4797 a
XX
PN US5636616-A.
XX
PD 17-JUN-1997.
XX
PF 10-NOV-1993: 94US-0152003.
XX
PR 18-NOV-1994: 94US-0442411.
XX
PR 10-NOV-1994: 94US-0152003.
XX
PA (ARCH-) ARCH DEV CORP.
XX
PI Liao S, Song C)
XX
DK WPI: 1997-432043/10.
XX
DB P-PSDP: W25034.
XX
PI DNA encoding ubiquitous nuclear receptor polypeptide(s) for
XX
PI producing recombinant polypeptide(s) and detecting RNA
XX
PS Claim 1: Column 67-70: 67bp: English.
XX
CC 179634 encodes a human ubiquitous nuclear receptor polypeptide (UR).
XX
CC The UR is useful in assays designed to select for a cell b
XX
CC interact with the UR which may potentially be of use therapeutically.
XX
CC UR is also used in diagnosis and drug design.
XX
SQ Sequence 1898 bp: 436 A: 594 C: 541 G: 222 T: 9 others
alignment_scores:
quality: 2923.00 length: 461
ratio: 4.748 gaps: 5
Percent Similarity: 92.625 Percent Identity: 86.551
alignment_block:
US-09-365-576-3 x 179634
Align seq 1/1 to: 179634 from: 1 to: 1898
1 MetSerSerProThrSerSortleuAspThrProValProclValAsnGlySer 17
|||||
71 ATGTCTACTGCCAAAGATTCCTGTGTATACCCGCTGACCTGGAATATATATCT 120

```

```

17 ProclProSerThrSerAlaThrSerProThrIleTylGlnGlnGly 44
|||||
121 GCTTCAAGATTCGGGAGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 170
44 Indle.....ThAspPro.....ProIleSerGlnGly 44
|||||
171 GGGAGGCTTGGGCGGAGGCTGGGAGGCTTGGGAGGCTTGGGAGGCTTGGGAG 220
45 SerSerSerAlaTylTleVal.....ValIleLeuGlnProclValAsp 59
|||||
221 GCTTCAAGATTCGGGAGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 270
59 ProclValIleValIleValIleValIleValIleValIleValIleVal 76
|||||
271 ACATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 320
76 ValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 92
|||||
321 AGCTTTCAGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 370
92 LeuSerTylGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 109
|||||
371 CTGACATCGGAGAGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420
109 GclValIleGlnValTylValIleValIleValIleValIleValIleVal 126
|||||
421 TGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 476
126 LARProMetAspMetAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 142
|||||
471 CTTCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 520
143 AlaGlnMetArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 159
|||||
521 GAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 570
159 SARGTleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 169
|||||
571 GAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 620
169 ProProProProSerGlnProAlaIleSerSerSerTylArgProAlaIle 185
|||||
621 CTTCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 667
186 SerProAlaThrSerSortleuAspThrProValProclValAsnGlySer 202
|||||
668 TGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 717
202 YlleGlnGlnThrAlaIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 219
|||||
718 TGTCTCAAGATTCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 767
219 LARProMetAspMetAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 235
|||||
768 GCGACATCGGAGAGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 817
235 ProclProSerThrSerAlaThrSerProThrIleTylGlnGlnGly 252
|||||
818 CCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 867
252 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 269
|||||
868 CTTCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 917
269 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 285
|||||
918 ACTGCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 967
286 IleAlaIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 302
|||||
968 ATGCGGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1017

```



```

|||||
691 CAGCTTAACTTAGAGTAACTTAAAGATATGATGAAAGTGGTGGTAACTCA 740
229 ntoultltyaenlysalqstvfthsseraspchlnfrolysvalthrtret 237
|||||
741 CCAACACAGCTAACAGACGCTGCTTTTCAGACGCTGCTGCTGCTGCTGCT 790
247 HPTROLGTYVALASPTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 253
|||||
791 GAGGCAATGCACTGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCT 840
254 AAlAsIthoPhthGlnGlnAlAlAlAlAlAlSerValGlnGlnAlAlAlAspPh 270
|||||
841 GCGCAATTTACTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 890
279 GAlAlAspGlnAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAl 287
|||||
891 GCGCAAAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 940
287 HAlGlnGlnAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAl 303
|||||
941 GCTTGTAAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 990
304 AAGTYSAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAl 320
|||||
991 AGATACAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1040
320 TserLysAspAspPhenISATGAlAlAlAlAlAlAlAlAlAlAlAlAlAl 337
|||||
1041 CAACTGAAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1090
347 TAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAl 353
|||||
1091 GATTTTTCATTTTCATTTTCATTTTCATTTTCATTTTCATTTTCATTTTC 1140
354 GAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAl 370
|||||
1141 GATTTTTCATTTTCATTTTCATTTTCATTTTCATTTTCATTTTCATTTTC 1190
370 GAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAl 387
|||||
1191 TGTGTAAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1240
387 TAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAl 403
|||||
1241 TCTTGTAAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1290
404 TProAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAl 420
|||||
1291 GAGGCAATTTACTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 1340
420 SserGlnGlnAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAl 437
|||||
1341 TTTCAAGAGTAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1390
447 GlnGlnSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 446
|||||
1391 TGGTGTAAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1419

```

seq_name: US09-365-576-3.rng seq_type: genomic_pos: 641-655 (6) size: 26

seq_documentation_block:

ID: Q39088 standard: cDNA: 1659 bp.

XX Q39088:

XX 24 JUN-1993 (first entry)

XX XR2 coding sequence.

XX Receptor: XR2: DNA binding domain: human; retinoid acid receptor-alpha;

XX hXR2-alpha; thyroid receptor-beta; vfk-beta; glucocorticoid receptor;

XX hXR2; retinoid X receptor-alpha; hXR2-alpha; vfk-beta; vfk-beta; vfk-beta;

XX hXR2; retinoid X receptor-alpha; hXR2-alpha; vfk-beta; vfk-beta; vfk-beta;

KW transcription-activation; response element; ss.

XX Homo sapiens.

XX Key location/Qualifiers

XX CDS 148..1470

XX /tag a

XX W09306215-A.

XX 01-Apr-1993.

XX 08-SEP-1992; 92W0-0807570.

XX 17-SEP-1991; 91US-0761068.

XX (SALK) SALK INST BIOLOGICAL STUDIES.

XX Borumeyer UK, Evans RM, Glanville V, Manolagiotis D;

XX Gng BS, Oro AE, Yao TP;

XX WPI: 1993-11756/14.

XX P-PSDB: R33744.

XX DNA encoding proteins of thyroid-steroid hormone receptor

XX superfamily - useful for screening for agonists-antagonists of

XX such receptors

XX Claim 9: Page 40-42: 71pp: English.

XX This sequence encodes the receptor XR2. The polypeptide encoded by

XX this sequence has a DNA binding domain comprising approx. 558 amino

XX acid sequence identity with the DNA binding domain of human retinoid

XX acid receptor-alpha (RAR-alpha), about 564 amino acid sequence

XX identity with the DNA binding domain of human thyroid receptor beta

XX (hTR-beta), about 508 amino acid sequence identity with the DNA

XX binding domain of human glucocorticoid receptor (hGR) and about 528

XX amino acid sequence identity with the DNA binding domain of human

XX retinoid X receptor-alpha (hRXR-alpha). XR2 receptor can be used for

XX testing a compound for its ability to regulate transcription

XX effects of a receptor polypeptide, identifying compounds which act as

XX ligands for the receptor polypeptides and for identifying response

XX elements for the receptor polypeptides.

XX Sequence 1659 BP; 499 A; 487 C; 454 G; 419 T; 0 other;

XX alignment_scores:

XX quality: 1485.00 length: 447

XX Ratio: 3.858 Gaps: 5

XX Percent Similarity: 80.313 Percent Identity: 62.864

XX alignment_block:

XX US-09-365-576-3 x Q39088 ..

XX Align seq 1/1 to: Q39088 from: 1 to: 1659

4 ProThrSerLeuAspThrProValProTyrAsnLysSerProGln 20

|||||

147 GCGCAATTTACTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 246

20 SerThrSerAlaThrSerProThrIleLysGlnGlnGlnGlnGlnGlnGln 47

|||||

237 GCGAGGCGGCAAGGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 284

37 SPTProProProGlnLysSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 54

|||||

284 TGGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 334

54 LoughLupProGlnAspGlnPro.....GlnAlAlAlAlAlAlAlAlAl 66

|||||

334 GAGAGGCGGCGCTTAGAGAGCGCAAGATATGCTGCTGCAAAAAAGGAGAAA 484

• • • •


```

TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-4000
TELEFAX: (714) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1959 base pairs
TYPE: nucleic acid
STRANDNESS: single
TOP-LOCATE: linear
FEATURES:
NAME/KEY: CDS
LOCATION: 256..1584
US: 08-442-411A-3

alignment_scores:
qualify: 2.247, 50
Ratio: 5.085
Percent Similarity: 99.104 Percent Identity: 97.085

alignment_block:
US: 09-365-576-3 x US: 08-442-411A-3
Align Seq 1/1 to: US-08-442-411A-3 From: 1 to: 1959

1 MetSerProThrSerLeuAspThrProValProGlyAsnGlySer 17
|||||
256 AAGTTTCCTCTTAAAGATTCTTGAAATCTCTCTGTGGGAATGGTTG 305
17 ProGlnProSerThrSerAlaThrSerProThrIleGlySerGlnGly 34
|||||
406 TCCGACGCTCACTACCTGCTCACTTACCCACTATTAAAGAAAGAGAG 455
44 InGlnThrAspProProGlnGlySerGlnGlySerSerAlaGlyIle 50
|||||
456 AAGACATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 405
51 ValValIleLeuGlnIleProGlnAspGlnProGlnGlySerGln 67
|||||
406 GTC.....GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 446
67 ValAlaProGlyMetLeuGlnIleSerGlnGlySerValGlySer 84
|||||
447 TGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 496
84 ValAlaSerGlnGlySerGlnGlySerValGlySerGlnGlySer 100
|||||
497 AAGCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 546
101 PhePheAlaGlnSerValValIleHisGlyValAlaGlyValGlyAlaGlySer 117
|||||
547 TTTCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 596
117 GlnGlySerGlnGlySerGlnGlySerValGlySerGlnGlySer 134
|||||
597 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 646
144 GlnGlySerGlnGlySerGlnGlySerValGlySerGlnGlySer 150
|||||
647 TTTCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 696
151 LeuSerGlnGlnGlnIleValAlaGlySerValGlnGlnGlnGlnGln 167
|||||
697 GTTCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 746
167 MetIleProProProSerGlnProAlaAlaSerSerSerGlyProAla 184
|||||
747 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 796
184 ValAlaSerGlnGlnGlnIleValAlaGlySerValGlnGlnGlnGln 200
|||||
797 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 846

```

```

201 GlnGlyIleGlnIleThrAlaAlaGlnGlnGlnGlnGlnGlnGln 217
|||||
847 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 896
217 ValAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 234
|||||
897 TGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 946
234 AlIleProIleProLeuGlnValAspProGlnSerAlaAspAlaGln 250
|||||
947 TCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 996
251 GlnAlaGlnAlaHisPheThrGlnGlnGlnGlnGlnGlnGlnGln 267
|||||
997 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1046
267 ValAlaPheAlaGlnGlnValProGlnGlnGlnGlnGlnGlnGln 284
|||||
1047 GGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1096
284 SerGlnIleAlaLeuLeuGlyAlaSerThrIleGlnIleMetLeuLeuGln 400
|||||
1097 ACCAGATGCGGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1146
401 ThrAlaArgArgGlyTyrAsnIleSerThrGlnGlnGlnGlnGln 417
|||||
1147 ACAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1196
417 PheGlnThrTyrSerTyrAspAspPheHisAlaGlnGlnGlnGlnGln 434
|||||
1197 GTTCACTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1246
434 IleIleAspProIlePheGlnIlePheSerAlaValIleAlaGlnGlnGln 450
|||||
1247 TCAATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1296
451 AspPheAlaGlnIleTyrAlaLeuLeuIleAlaIleAsnIleGlnSerAlaAs 467
|||||
1297 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1446
367 PheProAsnValGlnGlnProSerArgValGlnAlaLeuGlnGlnPro 484
|||||
1347 GCGGCTTAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1496
384 ValValGlnAlaLeuLeuSerTyrThrArgIleValAspProGlnIle 400
|||||
1397 ATGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1446
401 LeuAlaPheProArgMetLeuGlnGlnGlnGlnGlnGlnGlnGln 417
|||||
1447 GTGGCTTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1496
417 SerValHisSerGlnGlnGlnValPheAlaLeuArgGlnIleAspGlySer 434
|||||
1497 GTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1446
434 LeuProLeuLeuSerGlnIleTyrAspValHisGln 446
|||||
1547 TGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1584

```

seq_name: /cpn2_6/prodata/2/ina/pc/us_79MB_seq/p1-US94-1288-4

seq_documentation_block:

Sequence 4, Application pc/us9412883

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: CR10011005 NUTRIENT RICH FOOD COMPOSITIONS AND

TITLE OF INVENTION: METHODS

NUMBER OF SEQUENCES: 38

SEQUENCE ADDRESS:

ADDRESS: Arnold, White & Burke

STREET: P.O. Box 4433

[illegible]

```

Alignment: 100%
US-09-365-576-3 x US-08-649-6198-1 ...

Align seq 1/1 for US-08-649-6198-1 from: 1 to: 1979

1 MetSerPro...ThSerSerLaspThrProValProLysncl 16
|||||
206 ATGTGTTCTGCTACGAGAGATTGCTGTATACGCTGCTGTGAAATG 255

16 yserProInProSerThrSerAlaThrSerProLhrIleLysnclncl 43
1 |||||
256 CCGCTTTCAGGCTGAGCGGCTCTTCTTTCACCACTGTAAGAGAGAG 305

33 ycdlnh.....ThspproPro...proLysnclncl 43
|||||
306 GCGGCAAGCGTGAAGCGGCTGCGGCTGCGACGCTCATGCTCGAGGATGAT 355

44 GlySerSerSerAlaIyrIleVal.....ValIleLeuclnProLias 58
|||||
356 GAGGCTACGTACGCTGTACGCAACAGTCTGATTCATGCGATGCGTAAAG 405

58 pclurProLuarLysAqlLysAqlLysLysclProLualProLysMetLeucln 75
|||||
406 GGAAGTAAAGGCGCAAGCGTAAAGAAAGCGCGTCAAGCGCGCAACATCT 455

75 lscIleLysAqlYalIyScIYasLysLAsSerclYpMetlSerAsh 91
|||||
456 ACCAAGTTTGCGTGCTCTGTGGGAAAGAGCGCTCGGCTTCGACTAAAT 505

92 ValLAsSerCysScIleLysLysLysclYpMetPhePhePheValValH 108
|||||
506 GTCCTACAGTGGGAAAGCTGGCAAGCGCTTCTTCGCGGCGAGTGTGTCT 555

108 scIyScIyAlaIyAqlYalIyScIAsAqlLysSerclYpThrCysHlmetA 125
|||||
556 TGGTAAAGGCGCAAGCGCTATGCTTGGGAGGCTGGAGCAATCTGTCAAT 605

125 sPAlaPheMetLhrIleLysCysclLeLysPheAqlAqlATLysclYcl 141
|||||
606 ACCGCTTATGCGAGGCAAGTGGTAGAGTGGCTGTCTGGCAAGTGTAAAG 655

142 GluAlaIyMetArgclnclnLysValLeuSerclnclncllLeuAql 158
|||||
656 GAGCTAGAGATTAAGGAGACAGCTGCTCTTCTGAAAGAACATGCGGAA 705

158 sLysAqlLeuLysclnclnclnclnclnclnclnclnclnclnclncl 169
|||||
706 GAAAGAAATTCGGAAAGAGCAAGTAAAGAGATACATACACTGCTGACT 755

169 rclProLhrProSerclnclnProLualLAsSerSerSerclYalAqlProLual 185
|||||
756 CAGCTGTGGGCGCGTAAAGCGAGTAAAGTCAAGTCTGTGGGCTGTGGGCT 805

186 SerclYclYThrSerclnclnclnclnclnclnclnclnclnclncl 202
|||||
806 TCGCTGTGATCTGTGAAGTGAAGTGAAGTGAAGGCTGTGGGAGAGTGAAG 855

202 yIleclnclnclnAlaAlaIyPheLysMetlLeuclnclnclnclncl 219
|||||
856 TGTCTACCTAAAGTGGCTCAAAAGTAAAGATGACATGCTGTGGTGGCT 905

219 lscIleLuarLysAqlLysAqlYasLysCysPheSerSerclnclnclncl 235
|||||
906 CCGCACTGACAGTGGCAAAAGCGTTCGCTTCGCGAGCAATGGCAATCGAG 955

236 ProLhrProLuarLAsPProclnclnclnclnclnclnclnclnclncl 252
|||||
956 CCGCTGAGCGCTGGAGTGAAGACGCTCACTCGCGAATGTCTGGTCAAGTAA 1005

252 qPheLuarLAsPheLhrCysLualAqlIleLysValIyGlnclnclncl 269
|||||
1006 CTTTGCGCACTTCAGAGTGGGATATATCTATAGTAAAGATATGCTGT 1055

```


alignment_block:
 US-09-365-576-3 x US-08-646-248-1
 Alignment 1/1 to: US-08 646 248 1 from: 1 to: 2030
 Quality: 2033.50 Length: 463
 Ratio: 4.751 Gaps: 5
 Percent Similarity: 92.842 Percent Identity: 87.202
 1 Met-Ser-Pro...His-Ser-Leu-Asp-Thr-Pro-Val-Pro-Gly-Asn-1 16
 245 AGGCTCTCTCAACACGATTCTGTATACCGCCCTGGCTGGAATGG 294
 16 ysp-Pro-Glu-Pro-Ser-Ala-Ile-Ser-Pro-Thr-Lys-Ile-Ileu-33
 295 CCGCTTCAAGCTGGAGGCTCTTCTTCACCCACATGTAAGACAGAGAG 344
 33 Icy-Ile-Ile...His-Asp-Pro...Pro-Gly-Ser-Glu 43
 44 Gly-Ser-Ser-Ala-Ile-Ile-Val...Val-Ile-Ileu-Ile-Pro-Glu-58
 395 GAGGCAATCAAGCTTCTCAACATGGATCTGATCCAGATCCGAGAGA 444
 58 Pch-Ile-Pro-Ileu-Ile-Gly-Ser-Ile-Pro-Ileu-Ileu-Ileu-Ile 75
 445 GCAATCAACAGCTAAGCTAAGCAAGGCTGAGCGGCAACATGCTGAGCG 494
 75 Icy-Ile-Gly-Ser-Ala-Ile-Ile-Val...Val-Ile-Ileu-Ile-Pro-Glu-91
 495 ACAGATTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 544
 92 Val-Ileu-Ser-Gly-Ileu-Ileu-Ileu-Ileu-Ileu-Ileu-Ileu-Ile 108
 545 GCTCTCAATCTGCAAGGCTGCAAGGCTTCTTCTGCGATGCTATGCTGCG 594
 108 Ser-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile 135
 595 TCGTAAAGGCAAGCTGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 644
 135 Pch-Ile-Pro-Ileu-Ileu-Ileu-Ileu-Ileu-Ileu-Ileu-Ileu-Ile 141
 645 AGCTTTCTATGAGGCTGAGTCTCAAGTGGCTGGCTGGCTGGCTGGCT 694
 142 Glu-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile 158
 695 GAGGCAAGGATCAAGGATCAAGGATCTTCTTCTGCAACAGATCGGAGAA 744
 158 sly-Asn-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile 169
 745 GAAATCAATCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAG 794
 169 Pch-Ile-Pro-Ileu-Ileu-Ileu-Ileu-Ileu-Ileu-Ileu-Ileu-Ile 185
 795 CAACTGAGGAGCTGAGGATCAAGGATCAAGGATCTTGGGCTGGGAGCT 844
 186 Ser-Pro-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile 202
 845 TCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 894
 202 yll-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile 219
 895 TGTTCATCAATCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAG 944
 219 Icy-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile 235
 945 CCAATCTGAGCTAAGTAAAGCTCTTCTTCTGCAACAGTCAAGTCAAG 994
 240 Pro-Ile-Pro-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile 252
 995 CCGTAAAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 1044

252 Pch-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile 269
 1045 CTTTGGGCACTTCAAGGAGCTGAGGATCAATCAATCAATCAATCAATCA 1094
 269 sPh-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile 285
 1095 ACTGCTTAAGCAAGTGGCTGCTTCTCTGCTGCTGCTGCTGCTGCTGCT 1144
 286 ILe-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile 302
 1145 ATCGCCCTCTCGAAGGCAATCAATCAATCAATCAATCAATCAATCAATCA 1194
 302 GATGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 319
 1195 CAGCGGCTACAGACGACGACGACGACGATGATGATGATGATGATGATGAT 1244
 319 hTy-Ser-Ile-Asp-Asp-Phe-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile 345
 1245 CCTACAGCAAGGCAAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1294
 346 Asn-Pro-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile 392
 1295 AACCCATCTTGGACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1444
 352 PAla-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile 409
 1445 CCGTCAAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1494
 402 GPeP-Pro-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile 419
 1495 CTTCGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1544
 419 sHis-Ser-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile 445
 1545 TGCATCTGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1594
 446 Pch-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile-Ile 462
 1595 CCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1627
 seq_name: ysp-us-09-365-576-3-1
 seq_documentation_block:
 : Sequence 1, Application no. 7/99/6513924
 : GENERAL INFORMATION:
 : APPLICANT: Friedman, Ethan
 : APPLICANT: Holmway, M. Katharine
 : APPLICANT: Rodan, Gideon
 : APPLICANT: Rutledge, Su Jane
 : APPLICANT: Schmidt, Ariele
 : APPLICANT: Vogel, Robert
 : TITLE OF INVENTION: METHOD FOR FINDING RECEPTOR POTENTIALS
 : NUMBER OF SEQUENCES: 5
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Merck & Co., Inc.
 : STREET: 126 East Lincoln Avenue
 : CITY: Rahway
 : STATE: New Jersey
 : COUNTRY: US
 : ZIP: 07065-0907
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.25

IMMEDIATE SOURCE:
 CLONE: XP2 (XPL-SEC)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 148..1470
 US-08-433-358-7

alignment_scores:
 quality: 1495.00 length: 450
 ratio: 4.843 gaps: 7
 Percent Similarity: 80.667 Percent Identity: 63.111

alignment_block:

US-09-365-576-3 x US-08-433-358-7

Align seq 1/1 10: US-08 333-358-7 from 1 10: 1659

```

4 ProthrsrSerTouASPTrProValProgly...ASuclYserProcl 19
111 |-----|-----|-----|-----|-----|-----|
187 GCTGACTTGGGAGAGAGCTGCTGAGAGCCAGAGGACAGATGCAAGAGAG 246
19 nProthrsrSerTouASPTrProValProgly...ASuclYserProcl 36
111 |-----|-----|-----|-----|-----|-----|
247 GCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 282
46 hrASPProProProProProProProProProProProProProPro 50
111 |-----|-----|-----|-----|-----|-----|
283 .....ATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 424
51 ValValValValValValValValValValValValValValValVal 64
111 |-----|-----|-----|-----|-----|-----|
325 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 374
63 sPACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 80
111 |-----|-----|-----|-----|-----|-----|
475 GCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 424
80 aPACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 96
111 |-----|-----|-----|-----|-----|-----|
425 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 474
97 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 113
111 |-----|-----|-----|-----|-----|-----|
475 GCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 518
113 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 139
111 |-----|-----|-----|-----|-----|-----|
519 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 568
130 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 146
111 |-----|-----|-----|-----|-----|-----|
569 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 618
147 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 163
111 |-----|-----|-----|-----|-----|-----|
619 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 668
163 sPACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
111 |-----|-----|-----|-----|-----|-----|
669 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 718
180 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 196
111 |-----|-----|-----|-----|-----|-----|
719 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 738
197 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 213
111 |-----|-----|-----|-----|-----|-----|
749 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 767
213 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 230
111 |-----|-----|-----|-----|-----|-----|
768 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 817

```

```

230 sPACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 246
111 |-----|-----|-----|-----|-----|-----|
818 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 867
247 ASPAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 263
111 |-----|-----|-----|-----|-----|-----|
868 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 917
263 sPACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 280
111 |-----|-----|-----|-----|-----|-----|
918 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 967
280 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 296
111 |-----|-----|-----|-----|-----|-----|
968 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1017
297 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 313
111 |-----|-----|-----|-----|-----|-----|
1018 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1067
413 sPACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 430
111 |-----|-----|-----|-----|-----|-----|
1068 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1117
430 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 446
111 |-----|-----|-----|-----|-----|-----|
1118 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1167
447 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1183
111 |-----|-----|-----|-----|-----|-----|
1168 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1217
463 sPACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
111 |-----|-----|-----|-----|-----|-----|
1218 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1267
480 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 496
111 |-----|-----|-----|-----|-----|-----|
1268 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1317
497 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 513
111 |-----|-----|-----|-----|-----|-----|
1318 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1367
413 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 430
111 |-----|-----|-----|-----|-----|-----|
1368 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1417
430 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 446
111 |-----|-----|-----|-----|-----|-----|
1418 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1467

```

seq_documentation_block:

Sequence 7, Application 95/08463694

Patent No. 5696233

GENERAL INFORMATION:

APPLICANT: EVANS Ph.D., RONALD M.

APPLICANT: MANGELSDORF Ph.D., DAVID J.

APPLICANT: ONE MS., ESTELLITA S.

APPLICANT: ONE Ph.D., ANTHONY E.

APPLICANT: BOGMEYER Ph.D., OWE K.

APPLICANT: GIGHERE Ph.D., VICTOR MMN

APPLICANT: YAO M., ISO-PANG MMN

TITLE OF INVENTION: NEVIL RECEPTORS

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Preilly, Schroeder, Bruenemann & Clark

STREET: 444 So. Flower St., Suite 2000

CITY: Los Angeles

STATE: CA


```
413  UATGHIILeulSerSerValHisSerGluGlnValIleuAlaIleuArgLeuG 440
      |||||
1468  CCGGACCCGAGCCAGCGATCTCAGAGACAGAGGATGGTACCTGGGCTGGC 1417
      |||||
430  InAspLysLysLeuProIleuLeuSerGluIleuTrpAspValHisGlu 446
      |||||
1418  AGGACAAAAAGCTCCGACCGCTGGCTCTCTGACATCTGGATCTGGCACGAA 1467
```

-

-

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 29, 2001, 14:17:27 ; Search time 2390.16 seconds

(without alignments)
4792.079 Million cell updates/sec

Title: US-09-365-576-7

Percent score: 1860
Sequence: 1 GGCAGAGGACACAGAGTGGAA.....AAAAAAAAAAAAAAAAAAAA 1860

Scoring table: IDENTITY_ND2
Gapop 10.0 ; Gapext 1.0

Searched: 128325 seqs, 3078962356 residues

Total number of hits satisfying chosen parameters: 2566470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

listing first 45 summaries

```

GenBank:
1: qb_bal:*
2: qb_bal:*
3: qb_bal:*
4: qb_bal:*
5: qb_bal:*
6: qb_bal:*
7: qb_bal:*
8: qb_bal:*
9: qb_bal:*
10: qb_bal:*
11: qb_bal:*
12: qb_bal:*
13: qb_bal:*
14: qb_bal:*
15: qb_bal:*
16: qb_bal:*
17: qb_bal:*
18: qb_bal:*
19: qb_bal:*
20: qb_bal:*
21: qb_bal:*
22: qb_bal:*
23: qb_bal:*
24: qb_bal:*
25: qb_bal:*
26: qb_bal:*
27: qb_bal:*
28: qb_bal:*
29: qb_bal:*
30: qb_bal:*
31: qb_bal:*
32: qb_bal:*
33: qb_bal:*
34: qb_bal:*
35: qb_bal:*
36: qb_bal:*
37: qb_bal:*
38: qb_bal:*
39: qb_bal:*
40: qb_bal:*
41: qb_bal:*
42: qb_bal:*
43: qb_bal:*

```

```

44: em_ov:*
45: em_pat:*
46: em_ph:*
47: em_pl:*
48: em_ro:*
49: em_sts:*
50: em_sy:*
51: em_un:*
52: em_vl:*
53: qb_sts1:*
54: qb_sts2:*
55: qb_sts3:*
56: qb_sy:*
57: qb_sy:*
58: qb_vl1:*
59: qb_vl2:*
60: qb_vl3:*
61: qb_vl4:*
62: qb_vl5:*
63: qb_vl6:*
64: qb_vl7:*
65: qb_vl8:*
66: qb_vl9:*
67: qb_vl10:*
68: qb_vl11:*
69: qb_vl12:*
70: qb_vl13:*
71: qb_vl14:*
72: qb_vl15:*
73: qb_vl16:*
74: qb_vl17:*
75: qb_vl18:*
76: qb_vl19:*
77: qb_vl20:*
78: qb_vl21:*
79: qb_vl22:*
80: qb_vl23:*
81: qb_vl24:*
82: qb_vl25:*
83: qb_vl26:*
84: qb_vl27:*
85: qb_vl28:*
86: qb_vl29:*
87: qb_vl30:*
88: qb_vl31:*
89: qb_vl32:*
90: qb_vl33:*
91: qb_vl34:*
92: qb_vl35:*
93: qb_vl36:*
94: qb_vl37:*
95: qb_vl38:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SAMPLES

Result No.	Score	% Match	Query Length	ID	Description
1	1879.4	98.9	1841	94	MM000419
2	1589.2	85.4	1934	9	A49119
3	1569	84.4	1877	95	PM020394
4	1542.6	82.9	1959	10	146766
5	1542.6	82.9	1959	95	RM034375
6	1025	55.1	1898	10	146765
7	1023	55.0	1979	9	AK03516
8	1023	55.0	1979	45	E11456
9	1022.2	55.0	2030	10	136467
10	1022.2	55.0	2030	10	179211
11	1019.8	54.8	2010	93	HS067112

ALIGNMENTS

LOCUS	1841 bp	mRNA	F-0	03-MAR-1995
DEFINITION	Mus musculus retinoid X receptor interacting protein (RIP15) mRNA, complete cds.			
ACCESSION	U09419			
VERSION	009419.1			
KEYWORDS	house mouse; Mus musculus			
SOURCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eulipotyria; Rodentia; Sciurognathi; Muridae; Murinae; Mus; 1 (bases 1 to 1841)			
REFERENCE	Seol, W., Choi, H.S., and Moore, D.D.			
AUTHORS	Isolation of proteins that interact specifically with the retinoid X receptor: two novel orphan receptors			
TITLE	Mol. Endocrinol. 9 (1), 72-85 (1995)			
JOURNAL	95280959			
MEDLINE	2 (bases 1 to 1841)			
REFERENCE	Seol, W.			
AUTHORS	Direct Submission			
TITLE	Submitted (04-MAY-1994) Wongi Seol and David D. Moore, Dept. of Molecular Biology, Massachusetts General Hospital, Weisman W, Boston, MA 02114, USA			
FEATURES	Location/Qualifiers			
SOURCE	1..1841			
gene	Mus musculus "Mus musculus"			
gene	/db_xref="taxon:10090"			
gene	/db_xref="taxon:adult"			
gene	/tissue_type="liver"			
gene	131..1471			
gene	/gene="RIP15"			
gene	141..1471			
gene	/gene="RIP15"			

	query Match	Host Local Similarity	Matches 1840	Conservative	0	Mismatches	1	Indels	0	Gaps	
QY	1	60%	663	AA	A	A	A	A	A	A	1033
LB	1	60%	663	AA	A	A	A	A	A	A	1033
OY	61	60%	663	AA	A	A	A	A	A	A	1033
DB	61	60%	663	AA	A	A	A	A	A	A	1033
OY	121	60%	663	AA	A	A	A	A	A	A	1033
LB	121	60%	663	AA	A	A	A	A	A	A	1033
OY	181	60%	663	AA	A	A	A	A	A	A	1033
DB	181	60%	663	AA	A	A	A	A	A	A	1033
OY	241	60%	663	AA	A	A	A	A	A	A	1033
LB	241	60%	663	AA	A	A	A	A	A	A	1033
OY	301	60%	663	AA	A	A	A	A	A	A	1033
DB	301	60%	663	AA	A	A	A	A	A	A	1033
OY	361	60%	663	AA	A	A	A	A	A	A	1033
DB	361	60%	663	AA	A	A	A	A	A	A	1033
OY	421	60%	663	AA	A	A	A	A	A	A	1033
LB	421	60%	663	AA	A	A	A	A	A	A	1033
OY	481	60%	663	AA	A	A	A	A	A	A	1033
DB	481	60%	663	AA	A	A	A	A	A	A	1033
OY	541	60%	663	AA	A	A	A	A	A	A	1033
DB	541	60%	663	AA	A	A	A	A	A	A	1033
OY	601	60%	663	AA	A	A	A	A	A	A	1033
DB	601	60%	663	AA	A	A	A	A	A	A	1033

[illegible][illegible]

1b 114862 CATGACCTTGAGTGAAGAAATTAGGCTAGTTTCTCCCTGATGATTGGAGATTAAG 114921
QY 1781 GCGCTCTCTCTGCGGCTTTATTAATAAAAAATAAAATAAGTTTGAGTAATAGGTC 1840
114922 GCGCTCTCTCTGCGGCTTTATTAATAAAAAATAAGTTTGAGTAATAGGTC 114981
QY 1841 A 1841
1b 114982 A 114982

Search completed: April 29, 2001, 14:16:52
Job time: 7165 sec

•
•
•

GenCore version 4.5
Copyright (c) 1993 2000 Compaq Inc.

OM nucleotide: nucleotide search, using SW model

Run on: April 29, 2001 15:10:07 Search time for 6 records

(Without alignments)
7954.822 Million cell updates/sec

Title: US-09-365-576-7

Perfect score: 1860

Sequence: 1 GCCAGGCGCAACAGACGCGCA.....AAAAAAAAAAAAAAAAAAAA 1860

Scoring table: IDENTITY: 100
Gapop 10.0, Gapext 1.0

Searched: 678276 seqs, 201890651 residues

Total number of hits satisfying chosen parameters: 1556552

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database:

```
1: N_GenSeq_0401:
2: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
3: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
4: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
5: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
6: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
7: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
8: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
9: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
10: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
11: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
12: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
13: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
14: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
15: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
16: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
17: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
18: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
19: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
20: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
21: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
22: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
23: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
24: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
25: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
26: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
27: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
28: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
29: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
30: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
31: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
32: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
33: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
34: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
35: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
36: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
37: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
38: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
39: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
40: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
41: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
42: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
43: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
44: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
45: /SIFS2/seqdata/geneseq/geneseq/NA1981.DAT:
```

Prod. No. is the number of results produced by chosen to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1860	100.0	1860	17	Retinoid X receptor
2	1589.2	85.4	1934	17	OR 1 orphan receptor
3	1542.6	82.9	1959	16	Kat ubiquitination nuc
4	1542.6	82.9	1959	18	DNA encoding rat b
5	1025	55.1	1898	18	Human steroid recep
6	1022.2	55.0	2030	17	Human steroid recep
7	1022.2	55.0	2030	17	Human steroid recep
8	1021.4	54.9	1979	17	Human steroid recep
9	1020.6	54.9	2030	15	Human steroid recep
10	1012	54.4	1813	16	Human steroid recep
11	874.4	47.0	1622	21	Human steroid recep

ALIGNMENTS

RESULT 1	1	7954.8	38.5	1688	17	127617
1	1	526	28.3	1528	21	236043
2	1	522.8	28.1	1659	14	270098
3	1	517.4	27.8	1705	21	246040
4	1	494.6	26.5	1723	21	246042
5	1	432	23.2	1649	17	118554
6	1	152.2	8.2	2126	21	210412
7	1	147.6	7.9	2098	20	270408
8	1	147.6	7.9	2098	20	270408
9	1	147.6	7.9	2098	20	270408
10	1	147.6	7.9	2098	20	270408
11	1	147.6	7.9	2098	20	270408
12	1	147.6	7.9	2098	20	270408
13	1	147.6	7.9	2098	20	270408
14	1	147.6	7.9	2098	20	270408
15	1	147.6	7.9	2098	20	270408
16	1	147.6	7.9	2098	20	270408
17	1	147.6	7.9	2098	20	270408
18	1	147.6	7.9	2098	20	270408
19	1	147.6	7.9	2098	20	270408
20	1	147.6	7.9	2098	20	270408
21	1	147.6	7.9	2098	20	270408
22	1	147.6	7.9	2098	20	270408
23	1	147.6	7.9	2098	20	270408
24	1	147.6	7.9	2098	20	270408
25	1	147.6	7.9	2098	20	270408
26	1	147.6	7.9	2098	20	270408
27	1	147.6	7.9	2098	20	270408
28	1	147.6	7.9	2098	20	270408
29	1	147.6	7.9	2098	20	270408
30	1	147.6	7.9	2098	20	270408
31	1	147.6	7.9	2098	20	270408
32	1	147.6	7.9	2098	20	270408
33	1	147.6	7.9	2098	20	270408
34	1	147.6	7.9	2098	20	270408
35	1	147.6	7.9	2098	20	270408
36	1	147.6	7.9	2098	20	270408
37	1	147.6	7.9	2098	20	270408
38	1	147.6	7.9	2098	20	270408
39	1	147.6	7.9	2098	20	270408
40	1	147.6	7.9	2098	20	270408
41	1	147.6	7.9	2098	20	270408
42	1	147.6	7.9	2098	20	270408
43	1	147.6	7.9	2098	20	270408
44	1	147.6	7.9	2098	20	270408
45	1	147.6	7.9	2098	20	270408

Steroid hormone re
Human nuclear orph
X22 coding sequenc
Mouse CNP2B-1 cDNA
Rat orphan recepto
Retinoid acid rece
European corn bore
RAR-alpha gene, S
Human RAR-alpha cD
DNA of clone pRAR
myl/RAR-alpha fusi
Human pML/RARalpha
Human myl/RAR-alpha
pML-RAR-alpha RNA
Human pML/RAR-alpha
RAR-alpha-403 dupl
RAR-alpha clone 14
Helicobacter eclyson
Helicobacter eclyson
Synthetic modified
Bambyx mori nuclear
Synthetic modified
Modified eclyson
Retinoid X recepto
RAR gene, brossom
Clone pRAR-brossom
Retinoid X recepto
Fused human Reti
Human retinoid acti
Vector pRARX euro
Vector pRARX CA 5
Control vector pV4

Db	476	14tgcgcgctgctgcagagacgcctcgcggtctccactacaatggtctcaatttgtagg	535
Gy	420	gcctgcgaagcttctctccacagcagctgcggtctccactacaatggtctcaatttgtagg	479
Db	546	gcctgcgaagcttctccacagcagctgcggtctccactacaatggtctcaatttgtagg	595
Gy	480	gcctgcgaagcttctccacagcagctgcggtctccactacaatggtctcaatttgtagg	539
Db	546	gcctgcgaagcttctccacagcagctgcggtctccactacaatggtctcaatttgtagg	655
Gy	540	tcctgcgaagcttctccacagcagctgcggtctccactacaatggtctcaatttgtagg	599
Db	656	tcctgcgaagcttctccacagcagctgcggtctccactacaatggtctcaatttgtagg	715
Gy	600	gcctgcgaagcttctccacagcagctgcggtctccactacaatggtctcaatttgtagg	659
Db	716	gcctgcgaagcttctccacagcagctgcggtctccactacaatggtctcaatttgtagg	775
Gy	660	gcctgcgaagcttctccacagcagctgcggtctccactacaatggtctcaatttgtagg	719
Db	776	gcctgcgaagcttctccacagcagctgcggtctccactacaatggtctcaatttgtagg	835
Gy	720	gcctgcgaagcttctccacagcagctgcggtctccactacaatggtctcaatttgtagg	779
Db	846	gcctgcgaagcttctccacagcagctgcggtctccactacaatggtctcaatttgtagg	895
Gy	780	gcctgcgaagcttctccacagcagctgcggtctccactacaatggtctcaatttgtagg	839
Db	846	gcctgcgaagcttctccacagcagctgcggtctccactacaatggtctcaatttgtagg	905
Gy	840	gcctgcgaagcttctccacagcagctgcggtctccactacaatggtctcaatttgtagg	899
Db	956	gcctgcgaagcttctccacagcagctgcggtctccactacaatggtctcaatttgtagg	1015
Gy	900	gcctgcgaagcttctccacagcagctgcggtctccactacaatggtctcaatttgtagg	959
Db	1016	gcctgcgaagcttctccacagcagctgcggtctccactacaatggtctcaatttgtagg	1075
Gy	960	gcctgcgaagcttctccacagcagctgcggtctccactacaatggtctcaatttgtagg	1019
Db	1076	gcctgcgaagcttctccacagcagctgcggtctccactacaatggtctcaatttgtagg	1135
Gy	1020	gcctgcgaagcttctccacagcagctgcggtctccactacaatggtctcaatttgtagg	1079
Db	1146	gcctgcgaagcttctccacagcagctgcggtctccactacaatggtctcaatttgtagg	1195
Gy	1080	gcctgcgaagcttctccacagcagctgcggtctccactacaatggtctcaatttgtagg	1149
Db	1196	gcctgcgaagcttctccacagcagctgcggtctccactacaatggtctcaatttgtagg	1255
Gy	1140	gcctgcgaagcttctccacagcagctgcggtctccactacaatggtctcaatttgtagg	1199
Db	1256	gcctgcgaagcttctccacagcagctgcggtctccactacaatggtctcaatttgtagg	1315
Gy	1200	gcctgcgaagcttctccacagcagctgcggtctccactacaatggtctcaatttgtagg	1259
Db	1316	gcctgcgaagcttctccacagcagctgcggtctccactacaatggtctcaatttgtagg	1375
Gy	1260	gcctgcgaagcttctccacagcagctgcggtctccactacaatggtctcaatttgtagg	1319
Db	1476	gcctgcgaagcttctccacagcagctgcggtctccactacaatggtctcaatttgtagg	1435
Gy	1420	gcctgcgaagcttctccacagcagctgcggtctccactacaatggtctcaatttgtagg	1379
Db	1446	gcctgcgaagcttctccacagcagctgcggtctccactacaatggtctcaatttgtagg	1455
Gy	1440	gcctgcgaagcttctccacagcagctgcggtctccactacaatggtctcaatttgtagg	1459

Dd	1556	tgcgtctcgagatcaggatgtagataaatgaaggacccggacaactgcccacaccttgcatt	1941
Gy	1550	GGTGTCTTTCTTAAGACATGTCATCTTCAACTCTCTCTCTTGAACTTCCAGAGCATHTGCAC	1550
Dd	1616	TT	1616
Gy	1554	GCGCAGACACCGCCGCGGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTT	1614
Dd	1676	agcccaatccccccggcttcggatcagaagagagagagagagagagagagagagagagagagag	1728
Gy	1619	CCTACCATATTGATCTCTTCAAG	1676
Dd	1733	ccccaccgccgatgcagtcttccaag	1733
Gy	1679	CAGTAGCACCTCCACACACAGAGATCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG	1728
Dd	1793	cagctgcgccctctctcccgagctllaacctccagctccagctccagctccagctccagctccagct	1879
Gy	1739	AGACGCTTAAAGGCGAAGCTTGCGCGCGCGCGACACTTGCGGAGCGACAGACGCTTCTCTCTCT	1739
Dd	1853	aqaagttlaaqcgcaggttgctccaccatltgagagagagagagagagagagagagagagagag	1941
Gy	1799	TTATTATATAAAAAAACAAAATAATAGCT	1827
Dd	1913	tttatattaataaaaaataataataaat	1941
RESULT	5		
Tn	179634		
Tn	179634	standard; DNA; 1898 BP.	
XX			
Ac	179634;		
XX			
Dt	15-OCT-1997	(first entry)	
XX			
NF		DNA encoding human ubiquitous nuclear receptor polypeptide.	
KW		Ubiquitous nuclear receptor polypeptide; DNA binding domain;	
KW		therapy; drug design; diagnosis; ss.	
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	71..1453	
ET		/+lag-a	
XX			
PN	US5639616-A.		
XX			
PD	17-JUN-1997.		
PF	10-NOV-1993;	930S-0152003.	
XX			
PR	18-NOV-1994;	94US-0342411.	
PP	10-NOV-1994;	94US-0152003.	
XX			
FA	(ARCH-) ARCH NFV CAPP.		
XX			
P1	Liao S., Song G;		
XX			
WP1:	1997-332943/50.		
DR	P-PDBH; W25034.		
XX			
PT	DNA encoding ubiquitous nuclear receptor poly peptide(s) - for		
XX	producing recombinant poly peptide(s) and detecting RNA		
XX			
PS	Claim 1; Column 67-70; 67pp; English.		
XX			
CC	179634 encodes a human ubiquitous nuclear receptor polypeptide (UR).		
CC	The UR is useful in assays designed to select substances which		
CC	interact with the UR which may potentially be of use therapeutically		
CC	UR is also used in diagnosis and drug design.		
XX			
SO	Sequence 1898 BP; 436 A; 599 C; 541 G; 322 T; 0 other;		

QY	899	ATGAGCTAGCACTACAGCTGCGCGCTGAGAGAGATGTCAGATTGCTGCAAGCTGGCAGCG	958
DB	728	atgagctagcaactacagctgcgcgctgagagagatgtcagattgctgcaagctggcagcg	787
QY	959	TCTCTGACATTTGCGCGCGGAGAGAGAGATGCGCTGCTGCAAGCTGCTGCAAGATTGAGATC	1018
DB	788	tctctgacatcttgcgcgcgaggagagagatgcgctgctgcaagctgctgcaagattgagatc	847
QY	1019	ATGCTTCTTACAAATACAGCAAGCTTACAAATACAGCAAGCTTACAGCTTACAGCTTACAG	1078
DB	848	atgcttctttacaaatacagcaagcttacaaatacagcaagcttacagcttacagcttacag	907
QY	1079	GACTTCACTATACAGCAAGAGAGACTTCCAGCGTCGACAGCTTCGACAGTGGATATGATCAAT	1138
DB	908	gacttcactatcacagcaagagagacttccagcgtcgcacagcttcgacagtggaatgatcaat	967
QY	1139	CGTATCTTCGACATTTGCGCGCGGAGAGAGAGATGCGCTGCTGCAAGCTGCTGCAAGATTGCG	1198
DB	968	cgtatcttctgacatcttgcgcgcgaggagagagatgcgctgctgcaagctgctgcaagattgcg	1027
QY	1199	TGTGCTATATACATCAATCTCTTACAGCAAGTGGATGCTGAGAGAGAGAGAGAGAGAGAG	1258
DB	1028	tgtgctatatacatcaatctcttacagcaagtggatgctgagagagagagagagagagagag	1087
QY	1259	GTCAGACGCGCTGCGAG	1318
DB	1088	gtcagacgctgctgcag	1147
QY	1319	CGAAGCAAGCAAGCTTCCGCTTCCGACAGCATGCTTCATGACAGCTGCTGAGCGTGGCAGCTTC	1378
DB	1148	cgaagcaagcagcttgcgcttccgacagcatgcttcattgacagctgctgagcgtggcagcttc	1207
QY	1379	AGCTGCTCTTACAAATACAGCAAGCTTACAAATACAGCAAGCTTACAGCAAGCTTACAG	1438
DB	1208	agctgctctttacaaatacagcaagcttacaaatacagcaagcttacagcaagcttacag	1267
QY	1439	TTCGCTGCGAGAGATGCGAG	1498
DB	1268	ttcgctgctgcgagagatgctgcgagagagagagagagagagagagagagagagagagag	1327
QY	1499	CTGCT	1558
DB	1438	ctgct	1387
QY	1559	TGTGAG	1618
DB	1488	tgtgag	1447
QY	1601	GATATTAATACAGCTGCGAG	1660
DB	1508	gatattaatcacagctgctgcgagagagagagagagagagagagagagagagagagagag	1567
QY	1661	TCTGAG	1720
DB	1508	tctgag	1567
QY	1721	ATGCTTCTTACAAATACAGCAAGCTTACAAATACAGCAAGCTTACAGCTTACAGCTTACAG	1780
DB	1568	atgcttctttacaaatacagcaagcttacaaatacagcaagcttacagcttacagcttacag	1627
QY	1781	ATGCTTCTTACAAATACAGCAAGCTTACAAATACAGCAAGCTTACAGCTTACAGCTTACAG	1840
DB	1628	atgcttctttacaaatacagcaagcttacaaatacagcaagcttacagcttacagcttacag	1687

XX	Human nuclear orphan receptor LXR-alpha cDNA.
XX	
XX	Nuclear orphan receptor: LXR-alpha; human; rat; CNREB-1;
XX	CNRE binding factor; retin; negative regulatory element;
XX	cardiovascular disease; hypertension; congestive heart failure;
XX	diagnosis; therapy; transcription factor; ss.
XX	
XX	Homo sapiens.
XX	
XX	W09955345-A1.
XX	
XX	04-NOV-1999.
XX	
XX	23-APR-1999; 99WO-0508502.
XX	
XX	24-APR-1998; 98US-0082997.
XX	
XX	WPI: 2020-062002/05.
XX	
XX	Chen YE, Horinuchi M, Dzan VJ, Tamura K;
XX	
XX	(BGHM) BRIGHAM & WOMENS HOSPITAL INC.
XX	
XX	Claim 26; Page 77; 57pp; English.
XX	
XX	This is the nucleotide sequence of human nuclear orphan receptor
XX	LXR-alpha cDNA. It shows 49% homology to newly isolated cDNA (see
XX	US04) encoding mouse negative regulatory element binding factor
XX	CNREB-1 (see Y32374). CNREB-1 binds to the 5'-flanking region of
XX	the retin gene and modulates expression of the gene. CNREB-1 related
XX	polynucleotides, polypeptides and inhibitors are used in claimed
XX	methods for increasing or decreasing CNREB-1 activity in a
XX	mammalian cell or subject, e.g. for treatment of an adverse
XX	cardiovascular condition, hypertension, or congestive heart failure
XX	(claimed). In a claimed method, the CNREB-1 inhibitor is an
XX	antisense nucleic acid or a dominant negative nucleic acid of the
XX	present sequence. A claimed method for determining the level of
XX	CNREB-1 expression also utilizes the present sequence; an increase
XX	in CNREB-1 expression being indicative of an increased
XX	susceptibility of a subject for developing a retin-angiogenesis
XX	system mediated disorder such as hypertension or a cardiovascular
XX	disorder. LXR-alpha is a novel cAMP responsive transcriptional
XX	factor for human retin gene expression.
XX	
XX	Sequence 1528 BP; 340 A; 454 C; 439 G; 295 T; 0 other;
XX	
XX	Query Match 28.3%; Score 526; DB 21; Length 1528;
XX	Best Local Similarity 67.68, Fred No. 1,26 112;
XX	Matches 798; Conservative 0; Mismatches 325; Indels 57; Gaps 2;
XX	
XX	290 TTAAAGCCAAAGAGATGAGCTCAATGCTAAATGCTGAAGAAAGGCTGGCTGGAATATGCTT; 149
XX	
XX	255 tcagatccacacagatgacgctctcacaataaaggaagagcgacagccccaataatttctg 314
XX	350 GGGCATGAGCTGTGGCTGCTGATGGAGCAAGGCTTGAGGCTTCGATATGAAATGCTCT; 409
XX	
XX	315 gggaagacgacatcagcagctgtgttggagaaagctctggagcttcacataaatttttgc 474
XX	410 AGCTTAAAGCTGCAAAAGGCTCTCTGCTGGGAGCTATGCTGAGTGGGCTGGCTGGCT; 469
XX	
XX	375 agctggagagggctcaagaagattcttcggcagcgatcattcaa-----ggagagcgac 428
XX	470 TATGCTGTGTGGGGGCAAGGCAAGCTTGTGAGATGCAATGCTTATGCTATGCTGCTGCAAT; 529
XX	
XX	429 taccatctgcacacatgtggcgacactgccccatagacacatcactatcgtctgtatgtgcac 488
XX	530 CTCTGCGCGGCTCTGGCGAAGTGCAAAGAGGGTGGTATGGCGGAGCAATGCGTGTCTTGAG; 589
XX	

db 463 ggaagcagcctatgcttgcacacagcgttgatgcccacagatcctacatgagcgcg 522
uy 521 AAGTCCAACTCTGCGGAGCTGCGAAGTCCAAAGAGCTCGATGCGAAGTGGCTG 580
db 523 aaatgccagatcttgcacttcgcaatgccgcagcagcagatgaaagagagatgctg 582
uy 581 CTCTCTTAAATATATATTGGAAAGAAATATAGAAAGCAAGCAATATATAGCACTA 640
db 583 ctgtcagagagacatctcgtcttgaagatgaaagcagagagagagagagagagag 642
uy 641 CCGCCATCTTAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 700
db 643 ggcactctgctatccccaagagatccctacatctctcaatcctgctcctcctcctc 689
uy 701 GAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 760
db 690 -----acagctcagcctcaagagagctg 711
uy 761 CTGATGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 820
db 712 ggcataatcag 771
uy 821 CAGCCCAAGTAAAGCCCTGCGCCCTGCGCCCTGCGCCCTGCGCCCTGCGCCCTGCG 880
db 772 cgcctgagcctcagcctcagcctcagcctcagcctcagcctcagcctcagcctcag 831
uy 881 CAAGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 940
db 832 cagcctcagcctcagcctcagcctcagcctcagcctcagcctcagcctcagcctc 891
uy 941 GCTAAATGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 1000
db 892 gccaacacatctcctcagcctcctcagcctcagcctcagcctcagcctcagcctc 951
uy 1001 GCGTCGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 1060
db 952 acctctcagcctcagcctcagcctcagcctcagcctcagcctcagcctcagcctc 1011
uy 1061 TGATGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 1120
db 1012 agcctcagcctcagcctcagcctcagcctcagcctcagcctcagcctcagcctc 1071
uy 1121 CAGGTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 1180
db 1072 caggtgagatctcagcctcagcctcagcctcagcctcagcctcagcctcagcctc 1131
uy 1181 GACGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 1240
db 1132 aatgctcagcctcagcctcagcctcagcctcagcctcagcctcagcctcagcctc 1191
uy 1241 GTGATGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 1300
db 1192 gtcag 1251
uy 1301 TACAGGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 1360
db 1252 taagctcagcctcagcctcagcctcagcctcagcctcagcctcagcctcagcctc 1311
uy 1361 GTGATGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 1420
db 1312 ctgagcctcagcctcagcctcagcctcagcctcagcctcagcctcagcctcagcctc 1371
uy 1421 GACGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 1469
db 1372 gacaaag 1420


```

Db 1 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 60
QY 1 GAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
Db 61 GAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
QY 121 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 180
Db 121 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 180
QY 181 TCGTCGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 240
Db 181 TCGTCGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 240
QY 241 TCGTCGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 400
Db 241 TCGTCGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 400
QY 301 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 360
Db 301 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 360
QY 361 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 420
Db 361 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 420
QY 421 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 480
Db 421 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 480
QY 481 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 540
Db 481 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 540
QY 541 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 600
Db 541 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 600
QY 601 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 660
Db 601 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 660
QY 661 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 720
Db 661 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 720
QY 721 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 780
Db 721 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 780
QY 781 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 840
Db 781 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 840
QY 841 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 900
Db 841 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 900
QY 901 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 960
Db 901 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 960
QY 961 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 1020
Db 961 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 1020
QY 1021 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 1080
Db 1021 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 1080
QY 1081 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 1140
Db 1081 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 1140

```

```

QY 1141 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 1200
Db 1141 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 1200
QY 1201 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 1260
Db 1201 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 1260
QY 1261 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 1320
Db 1261 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 1320
QY 1321 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 1380
Db 1321 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 1380
QY 1381 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 1440
Db 1381 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 1440
QY 1441 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 1500
Db 1441 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 1500
QY 1501 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 1560
Db 1501 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 1560
QY 1561 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 1620
Db 1561 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 1620
QY 1621 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 1680
Db 1621 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 1680
QY 1681 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 1740
Db 1681 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 1740
QY 1741 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 1800
Db 1741 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 1800
QY 1801 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 1860
Db 1801 GCGAGGCTAACAGATCGAGATGCTGCGAGAGGCTGCGAGCTGCGAGCAT 1860

RESULT 2
PCT-US95-16311-7
Se-99-04-7, Application No. 7/99-04-7
GENERAL INFORMATION:
APPLICANT: Moore, David
APPLICANT: Seoul, Wong
APPLICANT: Choi, Hwang-Sik
TITLE OF INVENTION: RETINOID X RECEPTOR-INTERACTING
NUMBER OF SEQUENCES: 17
REFERENCE SEQUENCES:
ADDRESS: Fish & Richardson P.C.
STREET: 405 Franklin Street, Suite 1100
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2404
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

```


[illegible]

ADDRESSEE: Motec S Co., Inc.
STREET: 126 East Lincoln Avenue
CITY: Rahway
STATE: New Jersey
COUNTRY: US
ZIP: 07065-0907
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentis Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPROVED BY: 08/02/94
FILING DATE: 14-MAY-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/330,283
FILING DATE: 27-OCT-1994
APPROXY/AGENT INFORMATION:
NAME: Dolan, Catherine A.
REGISTRATION NUMBER: 36,502
REFERENCE/DOCKET NUMBER: 14427
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-4283
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2040 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYDROTHERMAL: NO
ANTI-SENSE: No
US-08-646-248-1

Db	1009	ATTCAGTGTGATGTTCTGTGAAATATCTGTGGAGGTAATAACTGTGGAGATGATCAATATCAGC	1068
Gy	1070	TTCTTAAAGCACTTCACTTATATGAAAGGACGATTTCAAGGCTGGAGGCTTGATAGTTGAAA	1129
Db	1069	TTCTTCAAGTATTTTCATTTTAATATGGGAAAGATTTGGTAAAGGTAAGGCTTGTAAATGTGAA	1128
Gy	1130	TTTCATCATCCCATCTTCGAAATTTCTGGGGGGCCATCGAGGCTGGGCTGGGCTGTGACATGCA	1189
Db	1129	TTTCATCATCCGCATCTTCGAATTTCTCCAGAGCCCATGAAATGAACTGGCAATTAATCATGTC	1188
Gy	1190	GAGTATGGCTTGGCTTATCGGCAATCAACATCTTCTCAAGGCGATGGGCTTAATATGTAGAGAG	1249
Db	1189	GAGTTTGGCTTGGCTATTTCTATTCAGATCTTCTGTGAGACCGGAGCAATCTGTAGAGAG	1248
Gy	1250	GTCACTCTTTGGAGGCTCTGTGATATGAGCTTATGTGGAGAGGCTCTCTCTTACAGGAGG	1309
Db	1249	GAGTGTGAGTGGAGGAGGCTCTGTGTATGATCAATATGTGGAGAGGCTCTCATATCTCTCCG	1308
Gy	1310	ATTCAGAGTGGGACAGGAGGATCTGGCTTGTGTACGATGATATCAAGCTGTGTGAGCTGTG	1369
Db	1309	ATTCAGCATGGCCATGAGGACATGATCTGTCCAGAGATGATATCAAAATATGTAAGCTCTG	1368
Gy	1370	CGCAGCCCTCAACTCTCGGTGTATCTGGAGGAGGCTCTTGTATTCGAGCTTCGAAATAGAGAG	1429
Db	1369	CGGAGCCCTGAGGAGGCTGTCAATATAGGCAAGATGTTGTGATCTGGCTGTGTATATATAAAAAG	1428
Gy	1430	CTGGGCTGCTTGTGCTGTGGAGATGTGGAGATGTGCATCAAGCT	1469
Db	1429	CTTCGATCTGCTGCTGTCTCGAATATGGGAGATGTGTACCAAT	1468

